

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:36:04 ; Search time 21.8726 Seconds
(without alignments)
955.720 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLQLSQPQSWTDD.....AEFYRLVSDVHGQSVWTP 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2580	100.0	504	1 TAB1_HUMAN	Q15750 homo sapien
2	190	7.4	300	1 P2C_PARTE	P49444 paramecium
3	175	6.8	399	1 P2C4_ARATH	P49598 arabidopsis
4	168	6.5	423	1 P2C2_ARATH	O04719 arabidopsis
5	167.5	6.5	434	1 P2C1_ARATH	P49597 arabidopsis
6	160.5	6.2	370	1 P2C3_SCHPO	Q09172 schizosacch
7	154	6.0	538	1 PDPI_RAT	O88483 rattus norv
8	151.5	5.9	491	1 P2C1_CABEL	P49595 caenorhabdi
9	148.5	5.8	356	1 P2C2_CABEL	P49596 caenorhabdi
10	147	5.7	529	1 PDPI_HUMAN	Q9p219 homo sapien
11	145	5.6	538	1 PDPI_BOVIN	P35816 bos taurus
12	143.5	5.6	414	1 P2C3_SCHPO	Q09173 schizosacch
13	143	5.5	538	1 PDPI_HUMAN	Q9p0j1 homo sapien
14	139.5	5.4	388	1 P2C3_ARATH	P49599 arabidopsis
15	134.5	5.2	383	1 P2C4_SCHPO	O14156 schizosacch
16	131.5	5.1	543	1 P2CG_BOVIN	P79126 bos taurus
17	128	5.0	382	1 P2CA_HUMAN	P35813 homo sapien
18	128	5.0	530	1 PDPI_RAT	O88484 rattus norv
19	127	4.9	382	1 P2CA_RABIT	P35814 oryctolagus
20	127	4.9	454	1 FEM2_HUMAN	P49593 homo sapien
21	126	4.9	382	1 P2CA_RAT	P20650 rattus norv
22	125.5	4.9	546	1 P2CG_HUMAN	O15355 homo sapien
23	125	4.8	382	1 P2C3_MOUSE	P49443 mus musculu
24	124	4.8	382	1 P2CA_BOVIN	O62829 bos taurus
25	122	4.7	347	1 P2C1_SCHPO	Q01371 schizosacch
26	121	4.7	2145	1 CYAA_PODAN	Q01513 podospora a
27	120	4.7	1692	1 CYAA_SCHPO	P14605 schizosacch
28	118	4.6	390	1 P2CG_MOUSE	P36993 mus musculu
29	118	4.6	542	1 P2CG_MOUSE	Q61074 mus musculu
30	118	4.6	2493	1 CYAA_USTMA	P49606 ustilago ma
31	117.5	4.6	393	1 P2C4_YEAST	P38089 saccharomyc
32	117	4.5	390	1 P2CB_RAT	P35815 rattus norv
33	117	4.5	406	1 P2C_LEICH	P36982 leishmania

RESULT 1

ID	TAB1_HUMAN	STANDARD;	PRT;	504 AA.
AC	Q15750;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	TAK1-binding protein 1.			
GN	TAB1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96216294; PubMed=8638164;			
RA	Shibuya H., Yamaguchi K., Shirakabe K., Tonegawa A., Gotoh Y.,			
RA	Ueno N., Irie K., Nishida E., Matsumoto K.;			
RT	"TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal			
RT	transduction.";			
RL	Science 272:1179-1182(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20057165; PubMed=10591208;			
RA	Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,			
RA	Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,			
RA	Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,			
RA	Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,			
RA	Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,			
RA	Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,			
RA	Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Goward M.E., Gratham D.V., Griffiths M.N.D., Hall C.,			
RA	Hall R.E., Hall-Ramlyn G., Heathcott R.W., Ho S., Holmes S.,			
RA	Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,			
RA	Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,			
RA	Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,			
RA	Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,			
RA	Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,			
RA	Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,			
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,			
RA	Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,			
RA	Vaadin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,			
RA	Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,			
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,			
RA	Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,			
RA	Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuoyama S.,			
RA	Roe B.A., Chen F., Chu Y., Hu P., Hua A., Kenton S., Do A., Do T.,			
RA	Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Do A., Do T.,			
RA	Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,			
RA	Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,			
RA	Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,			
RA	Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,			
RA	Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,			

ALIGNMENTS

ID	TAB1_HUMAN	STANDARD;	PRT;	504 AA.
AC	Q15750;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	TAK1-binding protein 1.			
GN	TAB1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96216294; PubMed=8638164;			
RA	Shibuya H., Yamaguchi K., Shirakabe K., Tonegawa A., Gotoh Y.,			
RA	Ueno N., Irie K., Nishida E., Matsumoto K.;			
RT	"TAB1: an activator of the TAK1 MAPKKK in TGF-beta signal			
RT	transduction.";			
RL	Science 272:1179-1182(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20057165; PubMed=10591208;			
RA	Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,			
RA	Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,			
RA	Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,			
RA	Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,			
RA	Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,			
RA	Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,			
RA	Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,			
RA	Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,			
RA	Gilbert J.G.R., Goward M.E., Gratham D.V., Griffiths M.N.D., Hall C.,			
RA	Hall R.E., Hall-Ramlyn G., Heathcott R.W., Ho S., Holmes S.,			
RA	Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,			
RA	Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,			
RA	Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,			
RA	Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,			
RA	Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,			
RA	Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,			
RA	Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,			
RA	Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,			
RA	Vaadin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,			
RA	Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,			
RA	Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,			
RA	Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,			
RA	Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuoyama S.,			
RA	Roe B.A., Chen F., Chu Y., Hu P., Hua A., Kenton S., Do A., Do T.,			
RA	Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Do A., Do T.,			
RA	Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,			
RA	Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,			
RA	Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,			
RA	Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,			
RA	Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,			

RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
RA Scher P., Walker C., Wamsley A., Wohldmann P., Peglin K., Nelson J.,
RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
RA McDermid H.E., Johnson A., Wong A.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dunanski J.P., Peyrard M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder K.E., O'Brien K.P., Lane L.,
RA Wilkerson P., Bodenreich A., Hartman K., Hu K., Khan A.S., Lane L.,
RA Tiliakou Y., Wright H.,
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
CC -!- FUNCTION: MAY BE AN IMPORTANT SIGNALING INTERMEDIATE BETWEEN TGFB
CC RECEPTORS AND MAP3K/TAK1.
CC -!- SUBUNIT: Interacts with MAP3K7 and with BIRC7.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS.
CC -!- SIMILARITY: CONTAINS 1 PP2C-LIKE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U49928; AAC12660.1; -;
DR EMBL: Z83845; CAB55304.1; -;
DR GenBank: HGNC:18157; MAP3K7IP1.
DR MIM: 602615; -;
DR InterPro: IPR001932; PP2C-like.
DR Pfam: PF00481; PP2C; 1.
DR SMART: SM00332; PP2C; 1.
FT DOMAIN 64 368 PP2C-LIKE.
FT SEQUENCE 452 457 POLY-SER.
SQ SEQUENCE 504 AA; 54644 MW; A45743288718983A CRC64;

Query Match 100.0%; Score 2580; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 3,1e-164;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MAQRSLLOSQSPMTDDLPCLSGSGASNRSGADGKGTSPEDSWLKFRSEN 60
1 MAQRSLLOSQSPMTDDLPCLSGSGASNRSGADGKGTSPEDSWLKFRSEN 60
61 NCFYGVENGVDGNRVTFVFAORLSAELLGQLNAEHADVRVLLQAFVVERSFLES 120
61 NCFYGVENGVDGNRVTFVFAORLSAELLGQLNAEHADVRVLLQAFVVERSFLES 120
121 IDDALEKASLSQSLPEGVPHQQLPROYOKILERLKTREISGAMVAVVILNNKLYV 180
121 IDDALEKASLSQSLPEGVPHQQLPROYOKILERLKTREISGAMVAVVILNNKLYV 180
181 ANVGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQLGLDAGKIKOYGIICQEST 240
181 ANVGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQLGLDAGKIKOYGIICQEST 240
181 ANVGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQLGLDAGKIKOYGIICQEST 240
241 RRIQYKVKYGYTIDILLSAASKPIIAPEIHGAQPLDGVTFVLVMSSEGLYKALEAAH 300
241 RRIQYKVKYGYTIDILLSAASKPIIAPEIHGAQPLDGVTFVLVMSSEGLYKALEAAH 300
241 RRIQYKVKYGYTIDILLSAASKPIIAPEIHGAQPLDGVTFVLVMSSEGLYKALEAAH 300
301 GFGQANQEIATAMITPEFAKQTSLDVAVAQVVDVRYRHSDFPAGSGEGARCPRHEDMTL 360
301 GFGQANQEIATAMITPEFAKQTSLDVAVAQVVDVRYRHSDFPAGSGEGARCPRHEDMTL 360
361 LVNRNGYPLGEMSOPTPSPAPAGRVYVPSVPSAOSTKSTSTLWVPSQGMWNG 420
361 LVNRNGYPLGEMSOPTPSPAPAGRVYVPSVPSAOSTKSTSTLWVPSQGMWNG 420
421 AHSASTLDEARPTLTNOSTNTHTOSSSSSSSDGGLFRSRPAHSLPAGEDGRVPEP 480
421 AHSASTLDEARPTLTNOSTNTHTOSSSSSSSDGGLFRSRPAHSLPAGEDGRVPEP 480

QY 481 YVDAEFYRLMSVDHGQSVVTAP 504
DB 481 YVDAEFYRLMSVDHGQSVVTAP 504

RESULT 2
P2C_PARTE
ID_P2C_PARTE STANDARD; PRT; 300 AA.
AC P49444;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE Protein phosphatase 2C (EC 3.1.3.16) (PP2C).
OS Parametium tetraurelia.
OC Eukaryota; Alveolata; Ciliophora; Oligohymenophorea; Periculida;
OC Parametium.
OX NCBI_TaxID=5888;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
AC STRAIN=51S;
RX MEDLINE=95105156; PubMed=7806499;
RA Klump S., Hanke C., Donella-Deana A., Beyer A., Kellner R.,
RA Plina L.A., Schultz J.E.;
RT "A membrane-bound protein phosphatase type 2C from Parametium
tetraurelia. Purification, characterization, and cloning.";
RL J. Biol. Chem. 269:32774-32780(1994).
CC -!- FUNCTION: ENZYME WITH A BROAD SPECIFICITY.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Membrane-bound.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z36985; CAB5448.1; -;
DR HSSP: P35813; IAG6.
DR InterPro: IPR001932; PP2C-like.
DR InterPro: IPR000222; PP2C.
DR Pfam: PF00481; PP2C; 1.
DR SMART: SM00331; PP2C StG; 1.
DR SMART: SM00332; PP2C; 1.
DR PROSITE: PS01032; PP2C; 1.
KW Hydrolyase; Magnesium; Manganese; Membrane.
FT METAL 37 37
FT METAL 38 38
FT METAL 57 57
FT METAL 237 237
FT METAL 289 289
SQ SEQUENCE 300 AA; 33739 MW; BC0318B4F7724EC CRC64;

Query Match 7.4%; Score 190; DB 1; Length 300;
Best Local Similarity 24.5%; Pred. No. 1e-05;
Matches 66; Conservative 51; Mismatches 104; Indels 48; Gaps 9;

64 LYGVENGVDGNRVTFVFAORLSAELLGQLNAEHADVRVLLQAFVVERSFLESID 123
64 LYGVENGVDGNRVTFVFAORLSAELLGQLNAEHADVRVLLQAFVVERSFLESID 123
52 VFGVFDHGGRVAVQFVEKHFDDELKKNK-----NFKEDQFEALKE 93
52 VFGVFDHGGRVAVQFVEKHFDDELKKNK-----NFKEDQFEALKE 93
124 ALAEKASLSQSLPEGVPHQQLPROYOKILERLKTREIS-GGAAVAVVAVVILNNKLYVAN 182
124 ALAEKASLSQSLPEGVPHQQLPROYOKILERLKTREIS-GGAAVAVVAVVILNNKLYVAN 182
94 TFLKDEL-----LTPEGCKELNQYKAVDTESYAGCTANVALLIKNTLYVAN 142
94 TFLKDEL-----LTPEGCKELNQYKAVDTESYAGCTANVALLIKNTLYVAN 142
183 VGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQLGLDAGKIKOYGIICQESTRR 242
183 VGNTRALLCKSTVDGLQVTLNDHTTENDELFRLSQLGLDAGKIKOYGIICQESTRR 242
143 AGDSRSVLCRNNTN-----HDMVDHKPNPEKSKITERAG---GVSDGRVAVGNINLSRA 195

RA Shpakovski G.V., Usery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: HAS AN IMPORTANT ROLE IN OSMOTIC STABILITY AND CELL
 CC SHAPE CONTROL. IT MAY NEGATIVELY REGULATE THE OSMOSENSING SIGNAL
 CC TRANSMITTED THROUGH WIS1 MAP KINASE.
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
 CC -1- SUBUNIT: MONOMER.
 CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L34881, AAA67320.1; -;
 CC EMBL: AL031579, CAA20880.1; -;
 CC DR HSPB, P35813, IABQ, PP2C-like.
 CC DR InterPro: IPR001932, PP2C-like.
 CC DR InterPro: IPR000222, PP2C.
 CC DR Pfam: PF00481, PP2C, 1.
 CC DR SMART: SM00332, PP2C SIG, 1.
 CC DR SMART: SM00332, PP2C, 1.
 CC DR PROSITE: PS01032, PP2C, 1.
 CC DR HydroLase, Magnesium, Manganese; Multigene family.
 CC KM METAL 37 37 MANGANESE 1 (BY SIMILARITY).
 CC FT METAL 38 38 MANGANESE 1 (BY SIMILARITY).
 CC FT METAL 63 63 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC FT METAL 233 233 MANGANESE 2 (BY SIMILARITY).
 CC FT METAL 282 282 MANGANESE 2 (BY SIMILARITY).
 CC SEQUENCE 370 AA; 40878 MW; DEDB826F44AC50 CRC64;
 SO
 Query Match 6.2%; Score 160.5; DB 1; Length 370;
 Best Local Similarity 20.7%; Pred. No. 0.0013;
 Matches 69; Conservative 63; Mismatches 105; Indels 97; Gaps 16;
 QY 51 DSWLKF-----RSENNCF-----YGVNGYDGNVTNFVQR 84
 DB 19 DRMLHFGVSHQGWRISEMDHACALNFTDSNSNPSTFGVGHGSDVAKYCRQHL 78
 QY 85 SAELLGQLNHAADVRVLLQAFDVERSFLSIDALAEKASLSQLEGVPOHQL 144
 DB 79 -PDITKSPSPFKKNYD-----EALKSGFL-ADNMLMODRDQ----- 115
 QY 145 PPQYKILERLKTLEREISGMAVAVVAVLNK-LYVANVGTNRALL-CKSTVDGLQVQ 202
 DB 116 -----EDPSGCTATTALIVDHQVIYCANAGSRVLGRKGTAE-----P 154
 QY 203 LNVHTTENEDLFRLSQLG--LDAGKIKQYGIIGQESTRIIGYKVKYVTDIDLSA 260
 DB 155 LSFHDKPNNDVEKARITAAAGFIDFGRVNA--GSLA---LSRLIGFEYK-----DSSLP 204
 QY 261 ASKRIPIAEPE--IHGAQPLDVGTFVLVMSGLYKALEAHAGPOAOEIAAMIDEFA 318
 DB 205 PEKQIVTAFPDVVHINIDPD---EFLIACDGIWDC-----KSSQVVEFVRGIV 253
 QY 319 KQTSLDAAVAQAVVDRVKRIHSDTPFASGGERARFC 352
 DB 254 AROSLEVICENIMDRCIASNSSESGICGDNMTIC 287
 RESULT 7
 ID PDP1 RAT STANDARD; PRT; 538 AA.
 AC 08483;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE [Pyruvate dehydrogenase [lipoamide]]-phosphatase 1, mitochondrial
 DE precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase,
 DE catalytic subunit 1) (PDP 1).
 GN PDP1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OK NCBI_taxid=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=98316337; PubMed=9651365;
 RA Huang B., Gudi R., Wu P., Harris R.A., Hamilton J., Popov K.M.;
 RT "Isoenzymes of pyruvate dehydrogenase phosphatase: DNA-derived amino
 RT acid sequences, expression, and regulation.";
 RL J. Biol. Chem. 273:17680-17688(1998).
 CC -1- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
 CC REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE
 CC PYRUVATE DEHYDROGENASE COMPLEX.
 CC -1- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
 CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
 CC -1- COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED.
 CC -1- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P4D PROTEIN OF
 CC UNKNOWN FUNCTION (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SKELETAL MUSCLE.
 CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AF062740, AAC40167.1; -;
 CC DR InterPro: IPR001932, PP2C-like.
 CC DR InterPro: IPR000222, PP2C.
 CC DR Pfam: PF00481, PP2C, 1.
 CC DR SMART: SM00331, PP2C SIG, 1.
 CC DR SMART: SM00332, PP2C, 1.
 CC DR PROSITE: PS01032, PP2C, 1.
 CC KM HydroLase, Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
 CC FT TRANSIT 1 71 MITOCHONDRION (BY SIMILARITY).
 CC FT CHAIN 72 538 [PYRUVATE DEHYDROGENASE [LIPAMIDE]]-
 CC SEQUENCE 538 AA; 61207 MW; D546096229B73564 CRC64;
 SO
 Query Match 6.0%; Score 154; DB 1; Length 538;
 Best Local Similarity 24.7%; Pred. No. 0.0067; Indels 140; Gaps 17;
 Matches 89; Conservative 38; Mismatches 93; Indels 140; Gaps 17;
 QY 40 DGKGTES-----HPPEDSWLKRSENNC-----FLYGVNGYDGNVTNFVQA 82
 DB 102 DGNVSSILGFDSNRNLPANAPIEDR---RSATTCIGTQGMILGVFDHAGACAGQAVSE 157
 QY 83 R-----LSAELLGQLNA----- 95
 DB 158 RLRYIAVSLPHEITLLEENAVESGRALLPILQMHKPNDFSKASKLYENGRLTYWQ 217
 QY 96 -----EHAEDVRRVLLQAFDVERSFLSIDALAEKASLSQLEGVPOHQLPQ 147
 DB 218 ELIDNTGESADIDVKEALINAF-----KRLDNDI-----SLEAQVGD-----PNS 258
 QY 148 YOKILERLKTLEREISGMAVAVVAVLNKLYVANVGTNRALLCKSTVDG-LQVQLNVND 206
 DB 259 FLNYL-----VLRVAFSGATACVAHVDGVDHVAANTGDSRAVLGVQEBDSASATVLSND 313
 QY 207 HTTENEDLFRLSQLDAGK-----IKQVGIIGQESTRIIGYKVKYVTDIDLS- 259
 DB 314 HNAQNERELQRLK---LEHPKNEAKSVVQDBRLILGLMFPFAFGVDFKFM---SIDLQKR 367


```

QY 89 LIGQLNHEADVRVILLQAFDVERSFLESTIDDLAKKASQSLQSGVPGHQLPPQY 148
DB 79 V---AOKESSEGMMK-----EALKEGPLE-----LDQOM-----104
QY 149 OKLERLKLTEREISGAMAVVAVLLNNKLYVANVTNRALLCKSTVDGLQVTLNVDT 208
DB 105 -----RVDEETADVSGTAAVVVLLKEGDIVYCGNAGDSRAV---SSVVG-EARPLSPFHK 155
QY 209 TENEDLFLPSQLGDAGKIKOVGIIICQES-TRRIGDYKVKYGTDTIDLSAASKPTI 267
DB 156 PSHEFEARRI-----IAAGWVEFNRRNGMLALSRALGDFAFK-----NCDTKP--199
QY 268 ASPEHIGAPLDGVNG-----FLVMSGLYKALBAHGGGQANOFIAMIDTEPAKQ 320
DB 200 AAEQIVTAF-P-DVITDKLTPDHEFTVLACDIWDVM-----TNQEVDPFVRKLAEK 250
QY 321 TSLDAVQAQVDR 333
DB 251 RDPQSIQCELLTR 263

RESULT 10
PDP2_HUMAN
ID _PDP2_HUMAN STANDARD; PRT; 529 AA.
AC Q9PZJ9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 2, mitochondrial
DE precursor (EC 3.1.3.43) (PDP 2) (Pyruvate dehydrogenase phosphatase,
DE catalytic subunit 2) (PDP 2).
GN PDP2 OR KIAA1348
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirogawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVI.
RT The complete sequences of 150 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:65-73(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
CC REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE
CC PYRUVATE DEHYDROGENASE COMPLEX (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC -!- COFACTOR: MAGNESIUM DEPENDENT (BY SIMILARITY).
CC -!- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P4D PROTEIN OF
CC UNKNOWN FUNCTION (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AB037769; AAA82586.1; ALT_INIT.
CC EMBL; BC028030; AAA828030.1; -.
CC InterPro; IPR001932; PP2C-like.
CC InterPro; IPR000222; PP2C.

```

```

DR Pfam: PF00481; PP2C; 2.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS01032; PP2C; 1.
KM Hydrolyase; Mitochondrion; Transit peptide; Magnesium.
FT TRANSIT 1 66 MITOCHONDRION (POTENTIAL).
FT CHAIN 67 529 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)] -
FT SEQUENCE 529 AA; 59978 MW; 252CAEBDAF61A5C CRC64;
SQ
Query Match 5.7%; Score 147; DB 1; Length 529;
Best Local Similarity 22.8%; Pred. No. 0.017;
Matches 84; Conservative 50; Mismatches 96; Indels 138; Gaps 17;
QY 45 ESHPEPDSWLKFRS-----ENNCLYGVNGYDGNRVTFVAAQL- 84
DB 99 ESRVP-NSVLRFSNOLAANSFVEDRRGVAACIQTGLMFGIFDGHGACAAQAVSERLF 157
QY 85 -SALLLQGLNHEA------DYRVLQAFDVERSTL 118
DB 158 YVAVSLMSHQTLHEMGAMESMKPLFILHMLKHPGDSIYKDVSVHLDLRLVYWQELL 217
QY 119 E-----SIDALA-----EKASIQSLPEGVPGHQLPPQYKILRLKLEREIS- 163
DB 218 DLMEMGLSIEELMYSFQRLDSDISLEIQAP-----LEDEVTR 256
QY 164 -----GGAMAVVAVLLNNKLYVANVTNRALLCKSTVDGL-QVTLNVDTTENED 214
DB 257 NLSLQVAFSGATCMAMHVGIIHIVANADCRALIGVEDNMGWCSCLPLTRDHNNANQAE 316
QY 215 LFLPSQLGIDA-----GKIQVGIIGQESTRIGDYKVKY-----GYTDI 255
DB 317 LSRKRHPESDRTTIMEDRLGVLP----RPFQVQLKSKELQSLTERGF-NT 370
QY 256 DLSAASKRP-----IIAPEI--HGAQPLDVGTVGLVMSGLYALBAHGGQA 305
DB 371 EALNIQFPPHYTPPYLTAPEVTVYHRLRPD---KLVLASDGLMML-----S 419
QY 306 NQETAMT 313
DB 420 NEDVRLV 427

RESULT 11
PDP1_BOVIN
ID PDP1_BOVIN STANDARD; PRT; 538 AA.
AC P35816;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 1, mitochondrial
DE precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase,
DE catalytic subunit 1) (PDP 1).
GN PDP1 OR PDP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxId=9913;
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
RX MEDLINE=93378941; PubMed=8396421;
RA Lawson J.E., Niu X.-D., Browning K.S., Trong H.L., Yan J., Reed L.J.;
RT "Molecular cloning and expression of the catalytic subunit of bovine
RT pyruvate dehydrogenase phosphatase and sequence similarity with
RT protein phosphatase 2C."
RL Biochemistry 32:8987-8993(1993).
CC -!- FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
CC REACTIVATION OF THE ALPHA SUBUNIT OF THE E1 COMPONENT OF THE
CC PYRUVATE DEHYDROGENASE COMPLEX.
CC -!- CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC -!- COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED.

```

```

CC -!- SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A PAD PROTEIN OF
CC UNKNOWN FUNCTION.
CC -!- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L18966; AAA30697.1; ALT_INIT.
CC PIR; A48692; A48692.
CC DR InterPro; IPR001932; PP2C-like.
CC DR InterPro; IPR00222; PP2C.
CC DR Pfam; PF00481; PP2C; 1.
CC DR SMART; SM00331; PP2C_SIG; 1.
CC DR SMART; SM00332; PP2CG; 1.
CC DR PROSITE; PS01032; PP2C; 1.
CC KW Hydrolase; Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
CC FT TRANSIT 1 71 MITOCHONDRION.
CC FT CHAIN 72 538 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)]-
CC FT PHOSPHATASE 1.
CC SQ SEQUENCE 538 AA; 61184 MW; 1A1C219AD8C3DAE3 CRC64;
Query Match 5.6%; Score 145; DB 1; Length 538;
Best Local Similarity 24.2%; Pred. No. 0.024;
Matches 87; Conservative 39; Mismatches 94; Indels 140; Gaps 17;
QY 40 DGKTES-----HPEDSWLKFRSENNC-----FLYFNGYDGNRTVTFVAQ 82
Db 102 DGNVSSVLGFDNSQLPANADIEDR----RSAATCLOTRGMLLGVPFGHAGCACSQAVSE 157
QY 83 R-----LSAELLGLQLNA-----
Db 158 RLFFVYVSLPHELTLEIENAVESGRALLPILQWKHPNDYFSEAKSLYFNSLRTYQ 217
QY 96 -----EHAEDVRVILLOAFDVVERSFLSDDDALAEKASLQSLPGVFOHQQLPPQ 147
Db 218 ELIDLTNGESTDIDVKEALINAF-----KRLNDI-----SLEAQVGD-----PNS 258
QY 148 YOKILERLKLTERISGAMAVAVLLNNKLYVANVTNRALLCKSTVDG-LQVTLQNVLD 206
Db 259 FLNYL-----VLRVAFSGATACVAHVDCVDLHVANTGDSRAMLGVSQEDGSWNAVTLSD 313
QY 207 HTTENEDFLRLSQLGLDAGK-----IKQVGIICGOESTRRIGDYKVYGYTDIDLLS- 259
Db 314 HNAQNEREVERLK---LSHPKNEAKSVVKQDRLLGLLMPFRAFGDVFKW---SIDLQKR 367
QY 260 AAKSKP-----IIAEPET--HCAQPLDGVTFVLVMSGLYKAL 296
Db 368 VTESGPDQLNDNEYTKFPPNYTPPYLTAPPEVTHRLRPQD---KFLVLTADGLMETM 424
RESULT 12
P2C3_SCHPO STANDARD; PRT; 414 AA.
AC Q09173;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Protein phosphatase 2C homolog 3 (EC 3.1.3.16) (PP2C-3).
GN PTC3 OR SPAC3G11.07C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;

```

```

RX MEDLINE=95163582; PubMed=7859738;
RA Shiozaki K., Russell P.;
RT "Counteractive roles of protein phosphatase 2C (PP2C) and a MAP
RT kinase kinase homolog in the osmoregulation of fission yeast.";
RL EMBL J. 14:492-502(1995).
RN [2]
RX SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21849401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Stevens K.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: HAS AN IMPORTANT ROLE IN OSMOTIC STABILITY AND CELL
CC SHAPE CONTROL. IT MAY NEGATIVELY REGULATE THE OSMOSENSING SIGNAL
CC TRANSMITTED THROUGH WIS1 MAP KINASE.
CC -!- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -!- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -!- SUBUNIT: MONOMER.
CC -!- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L34882; AAA67321.1; -.
CC DR EMBL; Z54354; CAA91172.1; -.
CC DR HSSP; P35813; IA6Q.
CC DR InterPro; IPR001932; PP2C-like.
CC DR InterPro; IPR00222; PP2C.
CC DR Pfam; PF00481; PP2C; 1.
CC DR SMART; SM00331; PP2C_SIG; 1.
CC DR SMART; SM00332; PP2CG; 1.
CC DR PROSITE; PS01032; PP2C; 1.
CC KW Hydrolase; Magnesium; Manganese; Multigene family.
CC FT METAL 37 37 MANGANESE 1 (BY SIMILARITY).
CC FT METAL 38 38 MANGANESE 1 (BY SIMILARITY).
CC FT METAL 62 62 MANGANESE 1 AND 2 (BY SIMILARITY).
CC FT METAL 230 230 MANGANESE 2 (BY SIMILARITY).
CC FT METAL 279 279 MANGANESE 2 (BY SIMILARITY).
CC FT CONFLICT 196 196 K -> T (IN REF. 1).
CC SQ SEQUENCE 414 AA; 44856 MW; EFP3A416625A2B11 CRC64;
Query Match 5.6%; Score 143.5; DB 1; Length 414;
Best Local Similarity 21.9%; Pred. No. 0.02;
Matches 62; Conservative 49; Mismatches 105; Indels 67; Gaps 12;
RC STRAIN=972;

```



```

QY 65 YGVFNGYDGNRYTNFYAQRSLAEELLGOLNABHADYRVLLQAFDVERSFILESIDA 124
Db 58 FAVYDGHGDYKAAKCGSNLPQIL---EKNPDPFGKDF-----VNALKSSFILN----- 102
QY 125 LAEKASLSQSLPEGVPOHQLPPOYOKILERLKTLEISGAMAVAVLNNKLYVANVG 184
Db 103 -ADKAILDD-----QFITDP-----SGCTATVVLVRGKLYCANAG 138
QY 185 TNRALCKSTVDGLOVTOQNVHTTENEDELFRLSQLG--LDAGKIKOVGIIICQESTRR 242
Db 139 DSRTVLGSKGI---AKPLSADHKPSNEAEKARICAGGFVDFGRVNG-----NLALSRA 189
QY 243 IGDYKVKYVYTDILLSAKSKPFIAPRPHGAQPLDVTGTVGLVMSGGLYKALEAAGP 302
Db 190 IDDFEFGKSNLE-----PEKQIVTALPQVVVHEITDD-DEFVVLACGDIWDC----- 235
QY 303 GOANOEIAAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASG 345
Db 236 -KTSQOVIEFVRGIVAGTSLERKIAENLMDNC--IASDIETTG 275

RESULT 13
PDP1_HUMAN STANDARD; PRT; 538 AA.
ID PDP1_HUMAN
AC Q9P0L1
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE [Pyruvate dehydrogenase (lipoamide)]-phosphatase 1, mitochondrial
DE precursor (EC 3.1.3.43) (PDP 1) (Pyruvate dehydrogenase phosphatase,
DE catalytic subunit 1) (PDP1C 1).
GN PDP1 OR PDP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP TISSUE=Adrenal gland;
RC MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Yu J.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Rong R., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhang C.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.";
RT Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
CC - FUNCTION: CATALYZES THE DEPHOSPHORYLATION AND CONCOMITANT
CC REACTIVATION OF THE ALPHA SUBUNIT OF THE EI COMPONENT OF THE
CC PYRUVATE DEHYDROGENASE COMPLEX (BY SIMILARITY).
CC - CATALYTIC ACTIVITY: [Pyruvate dehydrogenase (lipoamide)] phosphate
CC + H(2)O = [pyruvate dehydrogenase (lipoamide)] + phosphate.
CC - COFACTOR: MAGNESIUM DEPENDENT AND CALCIUM-STIMULATED (BY
CC SIMILARITY).
CC - SUBUNIT: HETERODIMER OF A CATALYTIC SUBUNIT AND A P4D PROTEIN OF
CC UNKNOWN FUNCTION (BY SIMILARITY).
CC - SUBCELLULAR LOCATION: Mitochondrial matrix (By similarity).
CC - SIMILARITY: BELONGS TO THE P2C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF155661; AAF67480.1; AUT_INIT.
DR InterPro; IPR001932; P2C-like.
DR InterPro; IPR000222; P2C.
DR Pfam; PF00481; P2C; 1.

```

```

DR SMART; SM00331; P2C; SIG: 1.
DR SMART; SM00332; P2CC; 1.
DR PROSITE; PS01032; P2C; 1.
KW Hydroxylase; Mitochondrion; Transit peptide; Magnesium; Calcium-binding.
FT TRANSIT 1 71 MITOCHONDRION (BY SIMILARITY).
FT CHAIN 72 538 [PYRUVATE DEHYDROGENASE (LIPOAMIDE)] -
FT PHOSPHATASE 1
SQ SEQUENCE 538 AA; 61219 MW; 1A3E512472540432 CRC64;
Query Match 5.5%; Score 143; DB 1; Length 538;
Best Local Similarity 24.7%; Pred. No. 0.032;
Matches 84; Conservative 38; Mismatches 90; Indels 128; Gaps 16;
QY 48 PPDSDMKFRSENNC-----FLYGVNNGYDGNRYTNFYAQRSLAEELLGOL 93
Db 122 PIEDR-----RSAATCLOTRGMILGVDFGHAGCAMSQAVSERLFYYIAGSLVPHETLLEIE 177
QY 94 NA-----EHAEDYRVLL 107
Db 178 NAYESGRALLPIQWKKHNDYFSKASLYRNSLRTYQELIDLTNGESTIDYKEALI 237
QY 108 QAFDVERSFILESIDALAEKASLSQSLPEGVPOHQLPPOYOKILERLKTLEISGAM 167
Db 238 NAF-----KRLDNDI-----SLBAQVGD-----PNSFLNYL-----VLRAFSGAT 273
QY 168 AVAVAVLNNKLYVANVTNRALLCKSTVNG-LQVTOQNVHTTENEDELFRLSQGLDNG 226
Db 274 ACVAHVDGVDLHVANTGDSRAMLVQGEEDGSWSAVTLSDHNAQNERELRLK--LEHP 330
QY 227 K-----IKOVGIICQESTRRIGDYKVKYGYTDIDLS-AAKSKP----- 265
Db 331 KSEAKSVKQDRLLGLMPFRAGDYKFKW---SIDLQKRVIESGPDQNDNEYTKFIP 387
QY 266 -----IIAEPI--HGAQPLDVTGTVGLVMSGGLYKAL 296
Db 388 NYHTPPLYTAEPEVTYHRLRPQD---KFLVLTATDGLWETM 424

RESULT 14
P2C3_ARATH STANDARD; PRT; 388 AA.
ID P2C3_ARATH
AC P49599; O9M0J6; Q9STP6;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein phosphatase 2C PPH1 (EC 3.1.3.16) (P2C).
GN PPH1 OR AT4G27800 OR T27E11.40.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=96026352; PubMed=7569999;
RA Schena M., Shalun D., Davis R.W., Brown P.O.;
RT "Quantitative monitoring of gene expression patterns with a
RT complementary DNA microarray.";
RL Science 270:467-470(1995).
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083488; PubMed=10617198;
RA Mayer K.F.X., Scheller C., Wambutt R., Entian K.-D., Terry N.,
RA Pohl T., Duesterhoeft A., Slickema W., Entian K.-D., Terry N.,
RA Harris B., Anesorge W., Brandt P., Grivell L.A., Rieger M.,
RA Weichselgartner M., de Simone V., Obermaier B., Meche R., Mueller M.,
RA Kreis M., Deleney M., Puligomenech P., Watson M., Schmidheini T.,
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
RA Vos P., Hobeisel J., Zimmermann W., Wedler H., Ridley P.,
RA Langham S.-A., McCullagh B., Bihlham L., Robben J.,
RA van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,

```

RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,
 RA Weitzneger T., Boche G., Ramsperger U., Hilbert H., Braun M.,
 RA Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 RA Moeljan P., Klein Lankhorst R., Rose M., Hauf J., Koeter P.,
 RA Berner S., Hempel S., Feldpausch M., Lamberth S., van den Daele H.,
 RA De Keyser A., Buyschaert C., Gielen J., Villaroel R., De Clercq R.,
 RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,
 RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McLay K., Mayes R.,
 RA Pettett A., Rajandream M.A., Lyne M., Benes V., Rechmann S.,
 RA Borkova D., Blocker H., Scharf M., Grimm M., Loehnert T.-H.,
 RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,
 RA Neumann S., Argirion A., Vitale D., Liguori R., Piravandi E.,
 RA Massenot O., Quigley F., Clabaud G., Muendlein A., Feiber R.,
 RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 RA Chedors F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,
 RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacón D., Jesse T.,
 RA Heijnen L., Schwarz S., Scholler P., Heber S., Frances P., Bielke C.,
 RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,
 RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,
 RA Parnell L., Dedhia N., Gnoj L., Schutz K., Huang E., Spiegel L.,
 RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 RA Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 RA Latrelle P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 RA Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 RA Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 RA Nelson J., Spieth J., Ryan E., Andrews S., Geisel C., Layman D.,
 RA Du H., Ali J., Berghoff A., Jones K., Drone K., Cotton M., Joshi C.,
 RA Antoniou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,
 RA Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R.,
 RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S.,
 RA Granat S., Shohdy N., Hasegawa A., Hameed A., Lochi M., Johnson A.,
 RA Chen E., Marra M., Martienssen R., McCombie W.R.,
 RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis
 RT thaliana";
 RL Nature 402:769-777(1999).
 CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
 CC phosphate.
 CC -1- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U34803; AAA92889.1; ..
 CC EMBL; AL161571; CAB81429.1; ..
 CC EMBL; AL078579; CAB43968.1; ALT_SEQ.
 CC InterPro; IPR001932; PP2C-like.
 CC InterPro; IPR000222; PP2C.
 CC Pfam; PF00481; PP2C; 1.
 CC SMART; SM00331; PP2C_SIG; 1.
 CC SMART; SM00332; PP2C; 1.
 CC PROSITE; PS01032; PP2C; 1.
 CC Hydrolase; Magnesium; Manganese; Multigene family.
 CC FT METAL 93 93 MANGANESE 1 AND 2 (BY SIMILARITY).
 CC FT METAL 296 296 MANGANESE 2 (BY SIMILARITY).
 CC FT METAL 326 326 MANGANESE 2 (BY SIMILARITY).
 CC FT CONFLICT 1 4 MALL -> SNSS (IN REF. 2).
 CC FT CONFLICT 69 70 RD -> QY (IN REF. 2).
 CC FT CONFLICT 271 271 D -> N (IN REF. 2).
 CC FT CONFLICT 353 353 W -> C (IN REF. 2).
 CC FT CONFLICT 360 388 RQNVVELVQAATTGLVTVGVMWSHLS -> TKRSR
 CC (IN REF. 2).
 CC SEQUENCE 388 AA; 42719 MW; 496ECC786AEB802 CRC64;
 SQ

Query Match 5.4%; Score 139.5; DB 1; Length 388;
 Best Local Similarity 21.3%; Pred. No. 0.034;

Matches 64; Conservative 58; Mismatches 99; Indels 79; Gaps 13;
 QY 63 FLY-GVFNQYDGNRVNFMVAQRLSAR-----LLLGQINAEHAEADVRRVLLQAFDVVERSF 117
 DB 86 FSYAAVFDGHAGSSSVKFLREELYKECVGALQAGSLNGGDFAAIKALIKAFESVDRNL 145
 QY 118 LESIDDAALAEKASLOSQLPQGVQHPQYQKILRLKTLREISGGAMAVVALLNKK 177
 DB 146 LKWL-----ANGDEESGSTATVIIRNDV 172
 QY 178 LVVANVTNRALLCKSTVDGLQVLTQNDVHTTENE-----DELFRLSQLGLDAGKIKQVG 232
 DB 173 SFTAHIGDSCAVLSRG---QIEELTDVHRPYGSSRAAIQVKKVKEAG---GWIVN-G 224
 QY 233 IICQGES-TRRTGDYKVKVGYTDI-----DILLSA--KSKPIIAEPETHGAP 277
 DB 225 RICGDIASVRAGDTRFTKKNMMLKKGVDGWRSEKTVSRIFKGMVMTPTDIFQV-P 283
 QY 278 LDGVTGFLVLMSEGLYKALEAAHGPQANQEAAMIDTEFAK-----QTSLDAAVAQVVD 333
 DB 284 LTSDFEIIILASDGLWDMYKSS-----DWSVVRDLRKHGNVQLACESLAQVALDR 335
 RESULT 15
 P2C4 SCHPO STANDARD; PRT; 383 AA.
 ID P2C4 SCHPO STANDARD; PRT; 383 AA.
 AC O14156; Q9UR02;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein phosphatase 2C homolog 4 (EC 3.1.3.16) (PP2C-4).
 GN PTC4 OR SPAC4A8.03C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A., SUBUNIT, SUBCELLULAR LOCATION, AND FUNCTION.
 RC STRAIN=972;
 RX MEDLINE=99365157; PubMed=10436019;
 RA Gaits F., Russell P.;
 RT "Vacuole fusion regulated by protein phosphatase 2C in fission
 RL [2].
 RN Mol. Biol. Cell 10:2647-2654 (1999).
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=211848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 RA Weltjens I., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;

```

RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: HAS A ROLE IN THE REGULATION OF VACUOLE FUSION.
CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein +
CC phosphate.
CC -1- COFACTOR: BINDS 2 MAGNESIUM OR MANGANESE IONS (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND. VACUOLAR.
CC -1- SIMILARITY: BELONGS TO THE PP2C FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF140285; AAD27651.1; -.
DR EMBL; Z98762; CAB58554.1; -.
DR InterPro; IPR001932; PP2C-like.
DR InterPro; IPR000222; PP2C.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR SMART; SM00332; PP2C; 1.
DR PROSITE; PS01032; PP2C; 1.
DR HydroLase; Magnesium; Manganese; Membrane; Multigene family.
KW METAL 92 92 MANGANESE 1 AND 2 (BY SIMILARITY).
FT METAL 308 308 MANGANESE 2 (BY SIMILARITY).
FT METAL 347 347 MANGANESE 2 (BY SIMILARITY).
SQ SEQUENCE 383 AA; 43569 MW; 88E00C2BDAE4B9BF CRC64;

Query Match 5.2%; Score 134.5; DB 1; Length 383;
Best Local Similarity 22.4%; Pred. No. 0.072;
Matches 59; Conservative 41; Mismatches 92; Indels 71; Gaps 10;

OY 63 FLYGVFNQYDGNRVNPFVAQRLSALLGQLNAEHAEDVRVLLQAFDV----- 112
DB 86 FFYGLFDHGGTECESEFLSTNGLKIIENQDLN-----DTEKILKEVHSVGVNAGLKP 139
OY 113 -VERSFLESIDDALEAKASLSQSLPEGVPOHQLPPQYOKILERLKTLEK-----EISGGAM 167
DB 140 FSLRTVLSGRDEDLMRARLYSF-----LQADMVYLNNVAPSPDSAVPGAV 187
OY 168 AVAVAVLNNK-----LVANVGTRALLCKSTVDGLQVTLNVHTTENEDEL 216
DB 188 GTVAITTSKNLSYWESESYIHLAVGDTALLCDSTRTG--RAHRLTFQHHPADVEAR 245
OY 217 RLSQGLDAGKIKQVGIICGF-----STRIGD-YKVKGYTIDILLSAKSKPIIAE 269
DB 246 RLRRYNNMFSRDS-----FGQRFAMVAVNTSFGGYKL-----KGLGVVAE 287
OY 270 PEIHGAQPLDGVTFVLIMSEGL 292
DB 288 PQLTSHSLRDMWFTLLSDGI 310

```

Search completed: December 9, 2002, 22:48:23
Job time : 23.8726 secs

Result No.	Query	Score	Match	Length	DB	ID	Description
1	2986	99.1	579	1	M3K7_MOUSE	Q62073	mus musculus
2	2982.5	99.0	606	1	M3K7_HUMAN	Q43318	homo sapien
3	501.5	16.6	393	1	M3K7_DROME	P83104	drosophila
4	493.5	16.4	954	1	M3K2_HUMAN	Q02779	homo sapien
5	418.5	13.9	394	1	M3K9_HUMAN	P80192	homo sapien
6	414	13.7	410	1	KYK2_DICD1	P18161	dictyosteli
7	412	13.7	859	1	M3KC_HUMAN	Q12852	homo sapien
8	408.5	13.6	888	1	M3KC_MOUSE	Q60700	mus musculus
9	408	13.5	898	1	M3KC_RAT	Q63796	rattus norv
10	398	13.2	821	1	CTRL_ARATH	Q05609	arabidopsis
11	379.5	12.6	1584	1	KYK1_DICD1	P18160	dictyosteli
12	373	12.4	832	1	ANR3_HUMAN	P57078	homo sapien
13	365.5	12.1	630	1	TEC_MOUSE	P42604	mus musculus
14	365.5	12.1	631	1	TEC_HUMAN	P24680	homo sapien
15	363	12.0	2347	1	KROS_HUMAN	P08922	homo sapien
16	362.5	12.0	1068	1	PAK1_XENLA	Q91738	xenopus lae
17	361	12.0	1055	1	PAK1_RAT	Q35346	rattus norv
18	358.5	11.9	620	1	ITK_HUMAN	Q08881	homo sapien
19	358.5	11.9	984	1	EP3A_RAT	Q08680	rattus norv
20	356.5	11.8	1052	1	PAK1_MOUSE	P34152	mus musculus
21	356	11.8	625	1	ITK_MOUSE	Q03526	mus musculus
22	354.5	11.8	527	1	TXK_MOUSE	P26682	mus musculus
23	354.5	11.8	993	1	EP3A_HUMAN	P29320	homo sapien
24	353.5	11.7	1114	1	RET_HUMAN	P07949	homo sapien
25	352	11.7	1115	1	RET_MOUSE	P35546	mus musculus
26	349.5	11.6	1052	1	PAK1_HUMAN	Q05397	homo sapien
27	349	11.6	380	1	KM1L_AVMH	P00531	avian retro
28	347.5	11.5	981	1	EP3A_BRARE	Q13146	brachydanio
29	346.5	11.5	804	1	FPS_DROME	P18106	drosophila
30	345.5	11.5	1620	1	ALK_HUMAN	Q9um73	homo sapien
31	344.5	11.4	983	1	EP3A_CHICK	P29318	gallus gall
32	344.5	11.4	1037	1	EP3A_HUMAN	P54756	homo sapien
33	343	11.4	467	1	MATK_RAT	P41243	rattus norv

SO SEQUENCE 579 AA; 64227 MW; 97C9F6F3C8E283EE CRC64;
Query Match 99.1%; Score 2986; DB 1; Length 579;
Best Local Similarity 99.1%; Pred. No. 9,6e-168;
Matches 574; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
DB 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
QY AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNVPCLVMEYAEAGSLYVNLHGAE 120
DB AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNVPCLVMEYAEAGSLYVNLHGAE 120
QY PLPYTAAAHMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGTVLKI CDFTAC 180
DB PLPYTAAAHMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGTVLKI CDFTAC 180
QY DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 181 DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVANGTRPPLINKLPKPIESLMTRCMSKDPQSRPMEIIVKIMTHLMRYFFGADEPLQY 300
DB 241 WAVANGTRPPLINKLPKPIESLMTRCMSKDPQSRPMEIIVKIMTHLMRYFFGADEPLQY 300
QY 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
DB 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATNGNQRRSIQDLTVYG 420
DB 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATNGNQRRSIQDLTVYG 420
QY 421 TEBGQVSSRSSSPSVAMITTSQPTSEKPTRSHPTPDDSTDNGSDNSIPMAYLTLDHQL 480
DB 421 TEBGQVSSRSSSPSVAMITTSQPTSEKPTRSHPTPDDSTDNGSDNSIPMAYLTLDHQL 480
QY 481 QPLAPCPNSKESMAVFEQHCMAQOEYKQVTEITALLQKQELVALLDQEDQNTSRL 540
DB 481 QPLAPCPNSKESMAVFEQHCMAQOEYKQVTEITALLQKQELVALLDQEDQNTSRL 540
QY 541 VOEHKKLIDENKSLSTYYQCKKOLEVIRSQQCKRGTS 579
DB 541 VOEHKKLIDENKSLSTYYQCKKOLEVIRSQQCKRGTS 579

RESULT 2
M3K7 HUMAN
ID _M3K7 HUMAN STANDARD; PRT: 606 AA.
AC O43318; O43317; O43319;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 15-JUN-2002 (Rel. 41; Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (BC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
GN MAP3K7 OR TAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN (1)
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC TISSUE=Lung;
RC MEDLINE=98153801; PubMed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.,
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-1 dependent mechanism",
RL Biochem. Biophys. Res. Commun. 243:545-549 (1998).
CC -I- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKs.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.

CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS, 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -I- SIMILARITY: IN THE C-TERMINAL SECTION, STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009357; BAA25026.1; -;
DR EMBL; AB009356; BAA25025.1; -;
DR EMBL; AB009358; BAA25027.2; -;
DR HSSP; P08631; IAD5
DR Genew; HGNC:6859; MAP3K7.
DR MTM; 602614; -;
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase.1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE; PSS0011; PROTEIN KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 8 14 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT BINDING 42 50 ATP (BY SIMILARITY).
FT ACT_SITE 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARSPIC 404 430 MISSING (IN ISOFORM 1A).
FT VARSPIC 509 518 PLAPCPNSKE -> ARTSCTGPG (IN ISOFORM 1C).
FT VARSPIC 519 606 MISSING (IN ISOFORM 1C).
SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;
Query Match 99.0%; Score 2982.5; DB 1; Length 606;
Best Local Similarity 95.4%; Pred. No. 1.6e-167;
Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;
QY 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
DB 1 MSTASAAASSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGRGAFGVCAKAKRAKOV 60
QY 61 AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNVPCLVMEYAEAGSLYVNLHGAE 120
DB 61 AIKOIESESEKKAFFIVELROLSRVNHNPVLYGACLNVPCLVMEYAEAGSLYVNLHGAE 120
QY 121 PLPYTAAAHMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGTVLKI CDFTAC 180
DB 121 PLPYTAAAHMSWCLQCSQGVAYLHSMOPKALIHRLKPPNLLVAGTVLKI CDFTAC 180
QY 181 DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIOTHTNNKSAAMMAPEVEGSSNYSEKCDVFSWGIIMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVANGTRPPLINKLPKPIESLMTRCMSKDPQSRPMEIIVKIMTHLMRYFFGADEPLQY 300
DB 241 WAVANGTRPPLINKLPKPIESLMTRCMSKDPQSRPMEIIVKIMTHLMRYFFGADEPLQY 300
QY 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
DB 301 PCOYSDGGSNSATSTGSEFMDIASTNTSNKSDTNMEQVATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATNGNQRRSIQDLTVYG 420
DB 361 SESGRSLGASHSSVESLPTSEGRKMSADMSIEFARIAATNGNQRRSIQDLTVYG 420

```

QY 404 -----GNGOPRRRSTODLTVTGTEPGQVSSSSSSPSVRMITTSPTSGTSEKPTRSH 453
Db 421 ILDPVEIVISGNGOPRRRSQDLTVTTEPGQVSSSSSSPSVRMITTSPTSGTSEKPTRSH 480
QY 454 WTPDDSTDTNGSNSIPMAYLTLDHQLQPLAPCNSKESMAVFEQCHKWAQYMKVOTEI 513
Db 481 WTPDDSTDTNGSNSIPMAYLTLDHQLQPLAPCNSKESMAVFEQCHKWAQYMKVOTEI 540
QY 514 ALLLQKQELVAELDQDEKQDQNTSLRVQEHKLLDENKSLSTYYQCKQLEIVRSQQ 573
Db 541 ALLLQKQELVAELDQDEKQDQNTSLRVQEHKLLDENKSLSTYYQCKQLEIVRSQQ 600
QY 574 KRQGTS 579
Db 601 KRQGTS 606

RESULT 3
M3K7_DROME STANDARD; PRT; 393 AA.
AC P83104;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Putative mitogen-activated protein kinase kinase kinase 7
DE (EC 2.7.1.-).
GN TAKL1.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.B.G.,
RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Feiler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Munkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."

```

```

RL Science 287:2185-2195 (2000).
RN [2]
RP CONCEPTUAL TRANSLATION.
RA Manning G., Sudarsanam S., Plowman G.;
RT "Prediction of novel protein kinases from the Drosophila genome
RT project and EST sequences.";
RL Unpublished observations (AUG-2001).
CC -!- FUNCTION: Can phosphorylate and activate yet undefined MAPKKS.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE003732; -; NOT ANNOTATED_CDS.
CC FlyBase; FBgn0046689; Takl1.
CC InterPro; IPR000719; Euk_Pkinase.
CC InterPro; IPR003290; Ser_Thr_Pkinase.
CC InterPro; IPR001245; Tyr_Pkinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00220; S_TKc; 1.
CC SMART; SM00219; TyRKc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00011; PROTEIN KINASE ST; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
KW Tyrosine-protein kinase; ATP-binding.
FT DOMAIN 11 266 PROTEIN KINASE.
FT NP_BIND 17 25 ATP (BY SIMILARITY).
FT BINDING 38 38 ATP (BY SIMILARITY).
FT ACT_SITE 133 133 BY SIMILARITY.
FT SEQUENCE 393 AA; 45238 MW; 43370C98490F0384 CRC64;
QY 29 EIDYKEIEV-EVVVGRGAFGVCKAKWRADVAIK-----QIESSEKAFIVELRQ 80
Db 3 KQVDFAEVKLSEKFLGAGSGAVRKATFQNOEIAVKIFDFLEETIKKNAER-----EITH 57
QY 81 LSRVNHPIVKLYGACLN--PVLVMEYAEGSLYLVHGAELPYPTAAAHAMSWCLQCS 138
Db 58 LSEIDHENVIRVIGRASNGKDYLLMEVLESGSLHNYLYGDDKWE-YTVEQAVRWALQCA 116
QY 139 QGVAYLHSMQPKALIHRLDKPNLLLVAGGTVLKICDPTGACDIQTHMNNKGSAAWMAP 198
Db 117 KALAYLHSLD-RPIVHRDIKQNMLLYNQHEDLKICDPLATDMNNKTKDMQGTLYMAP 175
QY 199 EVFEGSNSEKCDVPSGCIILWEVITRRKPFDEITGGP--AFRIMWVHNGTRPPL---IK 253
Db 176 EAIKHLKLYTAKDYYVFGIMLWELMTQLPYSHLENPNNSVAINKAISSGKLEPMEAVRS 235
QY 254 NLPKPIESILMTRCNSKDPSPSMBEIVKIMTHLMRVPPGADEPLQVPCQYSDGQSNSA 313
Db 236 DCFEGIKQLMECCMDINPEKPSMKIEKFLGE--QYESGTDDEFIKPL---DEDTVAUV 290
QY 314 T-----STGSFMDIASTNTSKSDTNMEQVPATNDTI-----KLESKLLKNQAKQOSE 362
Db 291 TVHVDSSGSR-----MRVDFWRHQLPSIRMTFPVVKREARELGLKTVVREMAKAAAD 342
QY 363 SGRSLGLSHSGSVESLPPTSEGGKMSADMSEIEARIAATGTNGQPPRRRSTQDLTVTQTE 422
Db 343 GDR-----EVRRAEKD-TERETSRAHNGERETRR-----AQGD 375
QY 423 PQQVSSRS 430

```

```

Db      376 VGETVRA 383

RESULT 4
M3KA_HUMAN STANDARD; PRT; 954 AA.
ID _M3KA_HUMAN
AC 002779; 012761; 014871;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 10 (EC 2.7.1.37)
DE (mixed lineage kinase 2) (Protein kinase MST).
GN MAP3K10 OR MEK2 OR MST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NC NCBI_TaxId=9606;
NX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96128179; PubMed=8536694;
RX Dorew D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
RA Sutherland G.R., Simpson R.J.;
RT "Complete nucleotide sequence, expression, and chromosomal
RT localisation of human mixed-lineage kinase 2.";
RL Eur. J. Biochem. 234:492-500(1995).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=95249256; PubMed=7731697;
RX MEDLINE=95249256; PubMed=7731697;
RA Katoh M., Hirai M., Sugimura T., Terada M.;
RT "Cloning and characterization of MST, a novel (putative)
RT serine/threonine kinase with SH3 domain.";
RL Oncogene 10:1447-1451(1995).
[3]
RN SEQUENCE OF 244-480 FROM N.A.
RP TISSUE=Colon epithelium;
RC MEDLINE=93238756; PubMed=8477742;
RX Dorew D.S., Devereux L., Dietzsch E., de Kretser T.;
RA "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; X90846; CAA62351.1; -.
CC EMBL; Z48615; CAA88531.1; -.
CC PIR; S32468; S32468.
CC HSRF; P11362; IFGK.
CC Genew; HGNC:6849; MAP3K10.
CC MIM; 600137; -.
CC InterPro; IPR000719; Euk_kinase.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR004040; STY_kinase.
CC InterPro; IPR002290; Ser_thr_kinase.
CC InterPro; IPR01245; Tyr_kinase.
CC Pfam; PF00018; SH3; 1.
CC Pfam; PF00069; kinase; 1.
CC PRINTS; PR00452; SH3DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_kinase; 1.

```

DR	ProDom:PD000066; SH3; 1.
DR <td>SMART; SMO0326; SH3; 1.</td>	SMART; SMO0326; SH3; 1.
DR <td>SMART; SMO0221; STYKC; 1.</td>	SMART; SMO0221; STYKC; 1.
DR <td>PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.</td>	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR <td>PROSITE; PS00108; PROTEIN_KINASE_ST; 1.</td>	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR <td>PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.</td>	PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR <td>PROSITE; PS50002; SH3; 1.</td>	PROSITE; PS50002; SH3; 1.
KW <td>Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; SH3 domain.</td>	Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase; ATP-binding; SH3 domain.
FT <td>DOMAIN 2</td>	DOMAIN 2
FT <td>DOMAIN 16</td>	DOMAIN 16
FT <td>DOMAIN 81</td>	DOMAIN 81
FT <td>NP_BIND 98</td>	NP_BIND 98
FT <td>NP_BIND 104</td>	NP_BIND 104
FT <td>BINDING 125</td>	BINDING 125
FT <td>ACT_SITE 222</td>	ACT_SITE 222
FT <td>DOMAIN 384</td>	DOMAIN 384
FT <td>DOMAIN 419</td>	DOMAIN 419
FT <td>DOMAIN 449</td>	DOMAIN 449
FT <td>DOMAIN 462</td>	DOMAIN 462
FT <td>CONFLICT 465</td>	CONFLICT 465
FT <td>CONFLICT 471</td>	CONFLICT 471
FT <td>CONFLICT 807</td>	CONFLICT 807
FT <td>CONFLICT 818</td>	CONFLICT 818
SO <td>SEQUENCE 954 AA; 103623 MW; 538f4AA559B0ABA CRC64;</td>	SEQUENCE 954 AA; 103623 MW; 538f4AA559B0ABA CRC64;
Query Match	16.4%; Score 493.5; DB 1; Length 954;
Best local Similarity	28.9%; Pred. No. 1.1e-21;
Matches 145; Conservative 82; Mismatches 192; Indels 83; Gaps 16;	
QY	21 APSQVLNFEEDIDYKEIEVEYVGRGAFGVCKAKRKADVAIKQIESSEKKAFIV----- 76
DB	83 APAGQLPQELPPEHQLQLEITIGVGFQKVRALRGEVAVKARLDPEKDPAATAGV 142
QY	77 --ELQSRVHNPVIVKLYGACLP--VCLWMEYAGESLVNVLHGAEPLPYTTAAHMS 132
DB	143 COEALLFGALQHPNIIALRGACINPPLCLVMEYARGALSRVLAGRVPVHV----LVN 198
QY	133 WCLQCSQGVAVYLSHMOFKALIHRLDKPNNLLV-----AGGYLKIICDFGTACD-IQT 184
DB	199 WAVYVARGMNYLHNDAPVPIIHRDLKSNILILBAIENHNLAADVLTIKITDGLAREMHKT 258
QY	165 HMTNNKGSAAWMAPEVFEGSNYSEKDVFSNGIIMVEYTRKRPPDEIGGPAFRIMAV- 243
DB	259 TKMSAAGTYAWMAPEVIRILSLFSKSDVWSFGVLLMELTGEVPREI--DALAVAYGVA 316
QY	244 HNGTAPPLIKLPEKIESLMTRCWKSQSPSPSMEIYKIMTHL-----MRYPQCA 294
DB	317 MNKLLPLPSTCFEPFALLLEECMDPDHGRDPDSILKRLVEIQSALFQWLPESFSL 376
QY	295 DEPLOYPCQY-----SDQGSNSATSTGSF-----MDIAST 325
DB	377 QEDWKLIEQHMFDDRLRTKEKELRSREELRLAAQQRQFQEBQLRRRQELARENDI--- 433
QY	326 NTSNKSDTNMEQVPAINTTIKRLSEKLLKNAQKQSESG--RLSLGSHGSVSESLPPT 382
DB	434 -VERELHLMQCQLSQEKRVRKRNKFNKRSRLKLREGSHSLSPSGEHNITVQASPTL 492
QY	383 SEGKMSADMSIEKRIATATGNGQRRRSIQDLVTGTEPGQVSSRRSSPSVNMITSG 442
DB	493 D--KKKSGDGAAPPSPSIT-----FLRLAIR--LTVVDCGSGSSGSSGSGSTWSPGG 542
QY	443 PTSEK-----PTRSHPWTPDD 459
DB	543 PPKEKELVGGKKKGKRTWGPST 564
RESULT 5	
TK9_HUMAN	STANDARD; PRT; 394 AA.
AC	P80192;
DT	01-FEB-1994 (Rel. 28, Created)
DT	01-FEB-1994 (Rel. 28, Last sequence update)


```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 9 (BC 2.7.1.1.-) (Mixed
DE lineage kinase 1) (Fragment).
GN MAP3K9 OR MLK1 OR PRKEL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC COLONIC BREAST AND OESOPHAGEAL ORIGIN.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- MAP KINASE KINASE SUBFAMILY.
DR PIR; S32467; S32467.
DR PIR; JU0229; JU0229.
DR HSSP; P12931; LFMK.
DR Gnew; HGNC:6861; MAP3K9.
DR MIM; 600136; -.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW ATP-binding.
FT NON_TER 1 1
FT DOMAIN 3 271 PROTEIN KINASE.
FT NP_BIND 9 17 ATP (BY SIMILARITY).
FT BINDING 30 30 ATP (BY SIMILARITY).
FT ACT_SITE 127 127 BY SIMILARITY.
FT DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
FT DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
FT DOMAIN 354 368 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;

Query Match 13.9%; Score 418.5; DB 1; Length 394;
Best Local Similarity 36.8%; Pred. No. 9.3e-18;
Matches 99; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

QY 35 EIEVEVVGAGFVCKAKWRAKDVAIK--QTESERKAFIVELRQ----LSRVNHPN 88
Db 2 ELTLEEIIIGGGKVTYRAFWGDEAVAKARHDPEDISQTIENVQAEKLFAMLKHPN 61
QY 89 IVKLYGACLN--VCLVMEYAEGLSNVLHGAPLPFYTTAAHAMSWCLQSCGVAYLHS 146
Db 62 IIALRGVCLKEPNLCLVMEFARGGFLNVLGSKRIPDI---LVNVAQIARGMYLHD 117
QY 147 MQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD-IQTHMTNKNKSAAMAP 198
Db 118 EAIVPPIIHRDLKSNLILQKVENGDLSNLIKLTIDTGLAREWHRTTQNSAAGTYAMAP 177
QY 199 EVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMAV-HNGTRPPLKNLPK 257
Db 178 EVTRASFSGKSDVMSVGVLLWELLTGEVPRGIDG--LRVAVGVAMNKLALPIPTCPE 235
QY 258 PIESLMTRCWSKDPSPQSPMEIEIVKIMT 285
Db 236 PFAKLMDOWNPDHPHSRPSFTNILDQLT 263

RESULT 6
KYK2_DICDI

```

```

ID KYK2_DICDI STANDARD; PRT; 410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.12) (Fragment).
GN PYK2 OR DPK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M33784; AAA33203.1; -.
DR PIR; B35670; B35670.
DR HSSP; P08631; 1AD5.
DR Dictyob; DD03011; DYKB.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT NON_TER 1 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B605B9AEC1 CRC64;

Query Match 13.7%; Score 414; DB 1; Length 410;
Best Local Similarity 31.5%; Pred. No. 1.8e-17;
Matches 104; Conservative 62; Mismatches 124; Indels 40; Gaps 12;

QY 28 FEE-----IDYKEIEVEEVVGAGFVCKAKWRAKDVAIKQI----ESESERK 72
Db 89 FEELKSIIGEREYIIDINDIQIKVGEGAFSEVWEGWKGIVHAIKLIIGDEQFKE 148
QY 73 AFIVELRSLRVNHPNIVKLYGACLNPLVCLVMEYAEGLSNVLH---GAPLPFYTTAAH 129
Db 149 RFIREVQNLKXGNHQNIVWFVIGACVKKPACIITEYMAGGLYNILHNPNSSTPKVYSFPL 208
QY 130 AMSWCLQSCQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKICDFTGACD--IQTHMT 187
Db 209 VLKMATDVALGLLHLHSI---TIVHRDLTSQNILLDELGNI-KISDFGLSAEKREGSMT 264
QY 188 NNKG---SAAWMAPEVPEG-SNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMAV 243
Db 265 MTNGGICNPRWRPPELTKNLGHYSEKVDVYCFSLVWWEILTGEIPFSDLDG-SQSAQVA 323
QY 244 HNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMEIEIVKIMTHLMRYFPGADEPLQYPCQ 303
Db 324 YAGLRPPIPEYCDPELKLTLQCEADPNDRPPFTYIVNKLKEI-----SWNNDIGF--- 375

```

QY 304 YSDEGQSNATSTGSEFMDIASTNTSKSDT 333
 DB 376 VSDQFYQYSEPTPR--LALSNQSSNSS 402

RESULT 7
 M3KC_HUMAN STANDARD; PRT; 859 AA.
 ID M3KC_HUMAN
 AC 012852;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK).
 GN MAP3K12 OR ZPK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Teratocarcinoma;
 RX MEDLINE=94311945; PubMed=8037767;
 RA Reddy U.R., Pleasure D.,
 RT "Cloning of a novel putative protein kinase having a leucine zipper
 domain from human brain.";
 RL Biochem. Biophys. Res. Commun. 202:613-620(1994).
 CC -1- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -1- COFACTOR: Magnesium.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -1- TISSUE SPECIFICITY: Highly expressed in brain and kidney.
 CC -1- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U07358; AAA67343.1; -
 CC HSSP: P12931; 1FMK.
 CC DR GeneW: HGNC:6851; MAP3K12.
 CC MIM: 600447; -
 CC DR InterPro: IPR000719; Euk_pkinase.
 CC DR InterPro: IPR004040; STY_pkinase.
 CC DR InterPro: IPR002290; Ser_thr_pkinase.
 CC DR Pfam: PF00069; pkinase; 1.
 CC DR ProDom: PD000001; Euk_pkinase; 1.
 CC DR SMART: SM00221; STYKC1.
 CC DR PROSITE: PS00107; PROTEIN_KINASE_ATP, FALSE_NEG.
 CC DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
 CC DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 CC DR Transferrase: Serine/threonine-protein kinase; ATP-binding;
 CC Phosphorylation; Magnesium; Membrane.
 CC FT DOMAIN 125 366 PROTEIN KINASE.
 CC FT NP_BIND 131 139 ATP (BY SIMILARITY).
 CC FT BINDING 152 152 ATP (BY SIMILARITY).
 CC FT ACT_SITE 236 236 BY SIMILARITY.
 CC FT DOMAIN 665 725 POLY-PRO.
 CC FT DOMAIN 720 725 POLY-GLU.
 CC SEQUENCE 859 AA; 93188 MW; 0E5209792C5C6F05 CRC64;

Query Match 13.7%; Score 412; DB 1; Length 859;
 Best Local Similarity 24.6%; Pred. No. 5.9e-17;
 Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEEVGARGAGVYCKAKMRAKDAVAIKQISESESRKAFIVELRQLSVNPN 89
 DB 119 EVPPPEILDLOWVSGAGQAGVFLGRFHGEYAVAKVADLKE-----IDIKLKLKHPNI 173
 QY 90 VKLYGACINP--VCLVMEYAGGSLVYVHLGAEPLPYRTAAHAMWSLQCSQGVAYLHSM 147
 DB 174 ITFGKVGTOGAPCCYCLMEFCAQGLYEVLRAGRPV---TSSLVDVMSGAGANNVYHLH 230
 QY 148 QPKLIRHDLKPPVLLVAGTGVKICDFGTACIQHMTNNK--GSAAMARVPEGSN 205
 DB 231 K---IHRDLKSPN-MLITVDVVKISDFGTSKLSDKSTKMSFAGTVAWMAPEVIRNEP 286
 QY 206 YSEKDVPSMGIIIMEVYTRRKPFDEIGPAPFRIMAV--HNGTRPPILKLPPIESLMT 264
 DB 287 VSEKVDIMSFQVLMELLTGELIPKVDSSA--IIVGSGNSLHLPPSSCPDGFILLR 344
 QY 265 RCWSKDPSPQSPSMEIIVKIMTHLMRPYPGADPLQPCQYSDGQSNATSTGSEFMDIAS 324
 DB 345 QCMNSKPRNRPSPFQ---ILHL-----DIAS 368
 QY 325 TNT-SNKSdT-----NMQVPATNDTIRLESKLL----- 353
 DB 369 ADVLSTPQETVFKSQAEWREVKLHFEKISEGTCLEHREELVMRRRELRHALDIRH 428
 QY 354 -----KNOAKQSSSGLR-----SLGASGSSVES 378
 DB 429 YERKLERANNLYMELNMLQLELKERELLREDALERRCPGLIKPPSGHLGNMTEK 488
 QY 379 L-----PTSEGRKMSADNSEIEPARI---AATGNGQPRRSIODLTVTGTEPGQVSS 428
 DB 489 LIKKRNVPQNLSPHSPQDPLKAEKSLPKLDALSLGVGLP-----GCPKAPSPGR--S 540
 QY 429 RSSSPSVMTTSGPTSEKP---TRSHPTPDDSTDING-----SUNSLPMAVLTIDHQ 479
 DB 541 RRGTRHRKSAKSCDPLRLTAVPHPHGGPSPGGLGGSPSAWEACPPALRGLHHD 600
 QY 480 L 480
 DB 601 L 601

RESULT 8
 M3KC_MOUSE STANDARD; PRT; 888 AA.
 ID M3KC_MOUSE
 AC Q60700; P70286;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
 DE (leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing
 DE kinase) (DLK).
 GN MAP3K12 OR ZPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Brain;
 RX MEDLINE=95074107; PubMed=7963011;
 RA Holzman L.B., Meritt S.R., Fan G.,
 RT "Identification, molecular cloning, and characterization of dual
 RT leucine zipper bearing kinase. A novel serine/threonine protein kinase
 RT that defines a second subfamily of mixed lineage kinases.";
 RL J. Biol. Chem. 269:30808-30817(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
 RX MEDLINE=96365386; PubMed=8769565;

RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
 RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
 RL DNA Cell Biol. 15:631-642(1996).
 RN [3]
 RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
 RX MEDLINE=96279269; PubMed=8663324; Holzman L.B.;
 RA Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
 RT "Characterization of dual leucine zipper-bearing kinase, a mixed
 RT lineage kinase present in synaptic terminals whose phosphorylation
 RT state is regulated by membrane depolarization via calcineurin.";
 RL J. Biol. Chem. 271:16888-16896(1996).
 CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -!- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
 CC testis, gastrointestinal tract, stomach, liver and pancreas.
 CC Within the nervous system, predominantly expressed in neurons and
 CC enriched in synaptic terminals.
 CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U14636; AA57280.1; --
 CC EMBL; U23789; AAB17123.1; --
 CC HSPSP; P12931; IFMK.
 DR MGD; MGI:1346881; Map3k12.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
 DR PROSITE; PS00108; PROTEIN KINASE-ST; 1.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding;
 FT Phosphorylation; Magnesium; Membrane.
 FT DOMAIN 158 399 PROTEIN KINASE.
 FT NP_BIND 164 172 ATP (BY SIMILARITY).
 FT BINDING 185 185 ATP.
 FT ACT_SITE 269 269 BY SIMILARITY.
 FT DOMAIN 56 62 POLY-GLY.
 FT DOMAIN 668 671 POLY-PRO.
 FT DOMAIN 698 701 POLY-GLU.
 FT DOMAIN 753 758 POLY-GLU.
 FT MUTAGEN 185 185 K->A: NO CATALYTIC ACTIVITY.
 FT MUTAGEN 192 192 E->A: NO CHANGE.
 FT CONFLICT 18 18 V->A (IN REF. 2).
 FT CONFLICT 28 29 KL->NV (IN REF. 2).
 FT CONFLICT 382 382 S->T (IN REF. 2).
 FT CONFLICT 494 495 EQ->DE (IN REF. 2).
 FT CONFLICT 517 517 N->D (IN REF. 2).
 FT CONFLICT 794 794 E->G (IN REF. 2).
 SQ SEQUENCE 888 AA; CFCEFD134F889ABB CRC64;
 Query Match 13.6%; Score 408.5; DB 1; Length 888;
 Best Local Similarity 24.9%; Pred. No. 9.9e-17;
 Matches 125; Conservative 76; Mismatches 157; Indels 145; Gaps 18;
 QY 30 EIDYKEIEVEVGRGAFGVVCKAKWRKADVAIKQIESESEKAFIVELRQLSRVNHPI 89

Db 152 EYPFEILLDQWVGSGAQAFLGRFHGEVAVKVRDLKE-----TDIKHLRLKLPENI 206
 QY 90 VKLYGACINP--VCLVMEYAGGSLYNVLHGAELPLPYPTAAHAMSWCLQCQGVAYLHSM 147
 Db 207 ITFGKVCTQAPCYCILMEFCAQQQLYEVLGRPV---TPSLLDVMSMGIAAGMNYLHLH 263
 QY 148 QPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKK--GSAAMWAPEVFECSN 205
 Db 264 K---IIHRDLKSPN-MLITYDDVKISDFGTSKELSDKSTKMSFAGTVAWMAPEVIRNEP 319
 QY 206 YSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLKPFIESLMT 264
 Db 320 VSEKVDIWSFGVVLWELLTGEIPYKDVDSA--IIVGVSNSLHLPVSSCPDGFKILLR 377
 QY 265 RCWSKDPQRORSMEIEIVKIMTHLMRYFFGADEPLQPCQYSDGQSNSATSTGSMFDIAS 324
 Db 378 QCWNSKPNRPFQFQ---ILLHL-----DIAS 401
 QY 325 TNT-SNKSDT-----NMEQVPATNDTIKRLESKLL----- 353
 Db 402 ADVLSTPQETVFKSQAEWRREVKKLHFEKIKSEGTCLHLEELVWRRREELRHLDIREH 461
 QY 354 -----KNOAKQOQSESGRL-----SLGASHGSSVES 378
 Db 462 VERKLERANNLYMELNMLQLKELKRELLRREQALERRCPGLLKSHPSRCLLHGNTMEK 521
 QY 379 L-----PPTSEGRKMSADMSEIARI-----AATGNGOPRRRSIQDLTVTGTSPGQVS- 427
 Db 522 LIKRNVQPKLSPHSKRPDIKTESLIPKLDALSGVGLP-----GCPKGPSP 570
 QY 428 --SRSSSPSVRMITTSPTSEKP 448
 Db 571 GSRRGKTRHKASAKSGCDLP 593
 RESULT 9
 M3KC RAT
 ID M3KC RAT STANDARD; PRT; 888 AA.
 AC Q63756;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DE 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 12 (BC 2.7.1.37)
 DE (MAPK-upstream kinase) (MUK).
 GN MAP3K12 OR MUK.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96226099; PubMed=8637721;
 RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
 RT "Activation of the JNK pathway by distantly related protein kinases,
 RT MEKK and MUK.";
 RL Oncogene 12:641-650(1996).
 CC -!- FUNCTION: May be an activator of the JNK/SAPK pathway.
 CC Phosphorylates beta-casein, histone 1 and myelin basic protein in
 CC vitro.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
 CC similarity).
 CC -!- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
 CC under basal conditions and dephosphorylated when membrane-
 CC associated (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; D49785; BAA08621.1; -.
DR HSSP; P12931; IFMK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN_KINASE
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP (BY SIMILARITY).
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
FT DOMAIN 753 758 POLY-GLU.
SQ SEQUENCE 888 AA; 96307 MW; 52AD964006BAE149 CRC64;

Query Match 13.5%; Score 408; DB 1; Length 888;

Best Local Similarity 28.3%; Pred.No. 1,1e-16; Mismatches 114; Indels 70; Gaps 12;

Matches 99; Conservative 67; Mismatches 114; Indels 70; Gaps 12;
QY 30 EIDYKEIEVEEVGRFAFVGCARARADVAIKQISESESEKAFIVEDRLSRVHPNI 89
DB 152 EVFEEBELDLQWVGSGAGQGVFLGRPHGEVAVKVRDLKE-----TDIKHLRKXKHPI 206
QY 90 VKLYGACLPN-VCLVMEVAGSLYNVLHGAEPLPYRAHAKMVCLOCSQGVAYLHSM 147
DB 207 ITRKGVCTQAPCYCILMEFCAQQLYEVLRAGRPV---TPSLVDVSMGIAGMNVLLH 263
QY 148 QPALKLHRDLKPNLLVAGCTYALKCDPCTACDICTHMTNNK--GSAAMMAPEVEGNS 205
DB 264 K---IHRDKSPN-MLITYDDVAKISDFTSKELSDKSTKNSFAGTAVMMAPVIRNRP 319
QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMAV-HNGTRPLINKLPKPIESLMT 264
DB 320 VSEKVDVMSGVVLMELLTGEIPIYKVDSSA--IIVGVSNSLHLVPSGCPGFKILR 377
QY 265 RCWSKDPSPQPSMEELVKTMTLMRYFPGADEPLQYPCOYSDGQNSATSTGSPNDIAS 324
DB 378 QCNMRKPRNRPFRQ---ILLHL-----DIAS 401
QY 325 TNF-SNKSDF-----NMEQVPATNDTIKRLESKLLKNQAKO 359
DB 402 ADVLSTPQETVFKSQAEWREEVKLHFEKISGECTLHRLLEELVMRRRE 451
RESULT 10
CTRL_ARATH
ID CTRL_ARATH STANDARD; PRT; 821 AA.
AC 005609;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase CTRL (EC 2.7.1.37).
GN CTRL OR AT5G03730 OR F17C15_150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=cv, Columbia; TISSUE=Seedling;
RX MEDLINE=33161417; PubMed=8431946;
RA Kleber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.,
RT "CTRL, a negative regulator of the ethylene response pathway in
RT Arabidopsis, encodes a member of the raf family of protein kinases";
RL Cell 72:427-441(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv, Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Konata M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,
RA Nakazaki N., Naito K., Okumura S., Shirao S., Takeuchi C., Wada T.,
RA Matanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozerky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchhoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ramsperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirke W., Moolman P., Klein Lankhorst R.,
RA Weitzenecker T., Bothe G., Rose M., Hauf J., Benneiser S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Meyer K.F.X., Rued S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
RT thaliana";
RL Nature 408:823-826(2000).
CC -1- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
CC PATHWAY.
CC -1- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -1- MISCELLANEOUS: CTRL MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC COTYLEDON GROWTH IS IMPAIRED.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC M1/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
DR EMBL; L08789; AAA32779.1; -.
DR EMBL; L08790; AAA32780.1; -.
DR EMBL; AL162506; CAB82938.1; -.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferase; Serine/threonine-protein kinase; ATP-binding;
FT DOMAIN 65 69 POLY-GLY.
FT DOMAIN 135 141 POLY-GLY.
FT DOMAIN 551 557 PROTEIN_KINASE.
FT NP_BIND 557 565 ATP (BY SIMILARITY).
FT BINDING 578 578 ATP (BY SIMILARITY).

DB 1561 ---EMEDQVSSFPAS 1572

RESULT 12
ANR3_HUMAN STANDARD; PRT; 832 AA.
ID ANR3_HUMAN STANDARD; PRT; 832 AA.
AC P57078; Q96KH0.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine-protein kinase ANKRD3 (EC 2.7.1.-) (Ankyrin repeat domain protein 3) (PKC-delta-interacting protein kinase).
GN ANKRD3 OR D1K.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal kidney, and Fetal lung;
RA Shimizu N., Kudo J., Shibuya K.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20289799; PubMed=10830953;
RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T., Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D., Reichwald K., Rump A., Schillabel M., Schudy A., Zimmermann W., Rosenthal A., Kudo J., Shibuya K., Kawasaki K., Asakawa S., Shintani A., Sasaki T., Nagamine K., Mitsuyma S., Antonarakis S.E., Minoshima S., Shimizu N., Nordiek G., Hornischer K., Brandt P., Schafte M., Schoen O., Desario A., Reichelt J., Kauer G., Blecker H., Ramser J., Beck A., Klages S., Hennig S., Resselmann L., Dagand E., Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F., Lehtach H., Reinhardt R., Yaspo M.-L.;
RT "The DNA sequence of human chromosome 21.";
RL Nature 405:311-319(2000).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are produced by alternative splicing.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 10 ANK REPEATS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-stb.ch/announce/> or send an email to license@isb-stb.ch).
CC -----
DB EMBL: AB047783; BAB56136.1; -
DB EMBL: AF001743; BAA95526.1; -
DB HSSP: P25963; I1KN.
DR Genew: HGNC:496; ANKRD3.
DR MIM: 605706; -
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000719; Euk_kinase.
DR InterPro: IPR004040; STY_kinase.
DR InterPro: IPR002290; Ser_thr_kinase.
DR InterPro: IPR001245; Tyr_kinase.
DR Pfam: PF00069; kinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_kinase; 1.
DR SMART: SM00248; ANK; 10.
DR SMART: SM00221; STYK; 1.
DR PROSITE: PS00107; PROTEIN KINASE_ATP; 1.
DR PROSITE: PS00108; PROTEIN KINASE_ST; 1.
DR PROSITE: PS00111; PROTEIN KINASE_DOM; 1.
DR PROSITE: PS00287; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00088; ANK_REPEAT; 9.

KW Transferase; Serine/threonine-protein kinase; ATP-binding; Repeat;
KW ANK repeat; Alternative splicing.
FT DOMAIN 22 286
FT REPEAT 485 514 ANK 1.
FT REPEAT 518 547 ANK 2.
FT REPEAT 551 580 ANK 3.
FT REPEAT 584 613 ANK 4.
FT REPEAT 617 647 ANK 5.
FT REPEAT 651 680 ANK 6.
FT REPEAT 684 713 ANK 7.
FT REPEAT 717 746 ANK 8.
FT REPEAT 750 780 ANK 9.
FT REPEAT 782 811 ANK 10.
FT NP_BIND 28 36
FT BINDING 51 51
FT ACT_SITE 143 143
FT VARSPIC 278 325
FT CONFLICT 714 714
SQ SEQUENCE 832 AA; 91610 MM; 5DBFFPD5F04F7ECB CRC64;
Query Match 12.4%; Score 373; DB 1; Length 832;
Best Local Similarity 27.9%; Pred. No. 1.1e-14;
Matches 168; Conservative 87; Mismatches 213; Indels 134; Gaps 35;

40 EVNGRGAFGVYCKAK---WAKQVAIKQIES-----ESERAFIVELRQLSRVNHPIVYL 92
DB EKVSQGGGQYQYKRVHWMWTW-LATCSPLHVDDEERELLEAKKMKAFRYLLPV 84
OY YGACINPCLVMEVYAEAGSYLVNLHGAELPY---YTAAMSMCLOCSQGVAVLSMOP 149
DB YGICREPVGLVMEVMEVSGLEKL-ASEPLPWDRPFIH-----ETAGNNFPHCMAP 137
OY 150 KALIHRIKIPENLLVAGTVYKICDFGA-CDIQH-----MTNKGSAAMNAP-VEE 202
DB -PLHLDLKIPANILLDAHYH-KISDFGLAKNGLSHDLSMDGLFGITAYLPPERIRE 195
OY 203 GSN-YSEKCVFSGIILMEVITRKPF-DEIGPARIMVAHNGR---PLIKLPLK 257
DB 196 KSRLEFDTHDVSPAIIVGWVLYQKKPFADK-N-IHIVKVVKGRRPLPVCRARPR 253
OY 258 P---IESIMTRCWGKDPSPSM-----EEIVKIMTHL---RY-FPG----- 293
DB 254 ACSHLIRLMQCMQGDPRVRRTPFGNGLNGLIRQVLAALLPVYGRRSRSGEGRLESEV 313
OY 294 ---ADEPLQYPCQVSDGQ-----SNSATSTGSEMDIANTSTNSKSDTNMEQVPATNDT 344
DB 314 IIRVTCPLSPQBITSETEDCEKPDDEVKETADLDVKS-----PEPRSEVVP---R 365
OY 345 IKRLSKLLKQNAQGESRGLSGA-----SHGSSVESLPTSGKMSAMSE 394
DB 366 LKRASAPTFDNDYSLSELSQLDSGVSQAVGPELRSRSSSESKLPSSGSGKRLSG-VSS 424
OY 395 IEARIATTTNGQ-----PRRSIQDLTVTGEBOGVSRRSSSPVIMTTSGPTSEPT 449
DB 425 VD---SAFSSRGSLSLSEPRRPSDICTDYGKKLV-----ALVSGTSLMK 472
OY 450 RSHPTPDSDTDTNGSDNSIPMAVLLTDHQLQPIAPC-----PNSKESMAVFEQCKM 502
DB 473 ILQPDVDLALDSGAS-----LHLAVEAGQEBEACAKMLLNANPNLSNRGSTPLH-M 525
OY 503 AOEYKVTETALLQKQELVAELDD-----EKQOQNSRLVQEHKKLLDENKS 553
DB 526 AVE-RRVGVVLELLARKISYNAK-DEDDWTALHFAAQNDESSTRL-----LLEKNAS 577
OY 554 LS 555
DB 578 VN 579
RESULT 13
TEC MOUSE
ID TEC MOUSE
AC P24604; STANDARD; PRT; 630 AA.


```

DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
GN TEC OR PSCTK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95019807; PubMed=7934162;
RA Sato K., Mano H., Ariyama T., Inazawa J., Yazaki Y., Hirai H.;
RT "Molecular cloning and analysis of the human Tec protein-tyrosine
  kinase.";
RU Leukemia 8:1663-1672(1994).
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
  tyrosine phosphate.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -I- TISSUE SPECIFICITY: HEMATOPOIETIC CELL LINES INCLUDING MYELOID,
  B-, AND T-CELL LINESAGES.
CC -I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
  SUBFAMILY.
CC -I- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -I- CAUTION: IT IS UNCERTAIN WHETHER MEN-1 IS THE INITIATOR.
CC -I- DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
  WWW="http://www.infobioenet.fr/services/chromocancer/genes/TECID75.html"
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use. By non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL, D29767; BAA06171.1; -.
DR HSBP; Q06187; IBS5.
DR Genew; HGNC:11719; TEC.
DR MIM; 600583; -.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_pkinase.
DR MIM; 600583; BTK; 1.
DR Pfam; PF00169; BTK; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00402; TECBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SMO0107; BTK; 1.
DR SMART; SMO0233; PH; 1.
DR SMART; SMO0252; SH2; 1.
DR SMART; SMO0326; SH3; 1.
DR SMART; SMO0219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH DOMAIN; 1.
KW Transferrase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation.
FT DOMAIN 4 111 PH.
FT DOMAIN 179 239 SH3.
FT DOMAIN 247 345 SH2.

```

FT	DOMAIN	370	623	PROTEIN KINASE.
FT	NP_BIND	376	384	ATP (BY SIMILARITY).
FT	BINDING	396	398	ATP (BY SIMILARITY).
FT	ACT_SITE	489	489	BY SIMILARITY.
FT	MOD_RES	519	519	PHOSPHORYLATION (AUTO-) (BY SIMILARITY)
SO	SEQUENCE	631 AA;	73629 MM;	A55DCARF991A9022 CRC64;
Query Match 12.1; Score 365.5; DB 1; Length 631;				
Best Local Similarity 34.7; Pred. No. 2.1e-14;				
Matches 94; Conservative 47; Mismatches 107; Indels 23; Gaps				
QY	30	EIDYEIEVEEYVGAGVWCKAKRAX-DVALIKOI-ESSEKRAFLVEIROLSRVNH	87	
DB	364	EINSEELFMRELGSGLFGVRLGKWRQYVAIKALREGAMCEEDFLFEAKVMKLTLP	422	
QY	88	NIIVKYGACL--NPCLVMEYAEGGSLYNVLHGAEPLPYTAAMAMSGCLQSGVAVLH	145	
DB	424	KLVOYLGVCTQOKPRLIYITFEMEGCLNLFRLQKQ--GHFRDVLVLSMCDVCEGMEL-	488	
QY	146	SMQPALIHRDLKPNLLLVAGTVLTKICDFGA---CDIQHTMTNNGSAAMAEVF	201	
DB	481	--ERNSEFHRDLAARNCLVSEAG--VVKVSDGMARYFLDDQYTSSGAKFPVKMCPEVF	53	
QY	202	EGSNSEKCDVFSNCILIMEVITR-RKPFEDIGPARIRIMAVANGR---PLIKLPLK	25	
DB	538	NYSRSSSDVWSFGVLMVEFTGGRPFEXYTN--YEVVTMTVRGRLYQPKLASNY--	593	
QY	258	PIESLMTRCMSKDPQSRPMEIKIMTHLM	288	
DB	594	-VVEYMLRCNOEKEBGRSPFDLRTIDELV	623	
RESULT 15				
KROS_HUMAN				
ID	KROS_HUMAN	STANDARD;	PRT;	2347 AA.
AC	P08922; 015368;			
DT	01-NOV-1968 (Rel. 09, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Proto-oncogene tyrosine-protein kinase ROS precursor (EC 2.7.1.112) (c-ros-1).			
GN	ROS1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90280463; Pubmed=2352949;			
RA	Bitchmeier C., O'Neill K., Riggs M., Wigler M.;			
RT	"Characterization of ROS1 cDNA from a human glioblastoma cell line.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 87:4799-4803(1990).			
RP	[2]			
RP	SEQUENCE OF 1790-2259 FROM N.A.			
RX	MEDLINE=87064611; Pubmed=3023956;			
RA	Matsushime H., Wang L.-H., Shibuya M.;			
RT	"Human c-ros-1 gene homologous to the v-ros sequence of ur2 sarcoma			
RL	MoJ. Cell. Biol. 6:3000-3004(1986).			
RN	[3]			
RP	SEQUENCE OF 1854-2245 FROM N.A.			
RX	MEDLINE=87064625; Pubmed=3785223;			
RA	Bitchmeier C., Birnbaum D., Matches G., Fasano O., Wigler M.;			
RT	"Characterization of an activated human ros gene.";			
RL	MoJ. Cell. Biol. 6:3109-3116(1986).			
CC	-1- FUNCTION: THIS IS A PROBABLY A CELL GROWTH OR DIFFERENTIATION			
CC	-2- FACTOR RECEPTOR WITH A TYROSINE-PROTEIN KINASE ACTIVITY.			
CC	-3- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein			
CC	-4- tyrosine phosphate.			
CC	-5- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-6- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. INSULIN			
CC	RECEPTOR SUBFAMILY.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M3453; AAA60278.1; -;
 DR EMBL; M1359; AAA60277.1; -;
 DR EMBL; M13368; AAA60277.1; JOINED.
 DR EMBL; M13591; AAA60277.1; JOINED.
 DR EMBL; M13592; AAA60277.1; JOINED.
 DR EMBL; M13593; AAA60277.1; JOINED.
 DR EMBL; M13594; AAA60277.1; JOINED.
 DR EMBL; M13595; AAA60277.1; JOINED.
 DR EMBL; M13596; AAA60277.1; JOINED.
 DR EMBL; M13597; AAA60277.1; JOINED.
 DR EMBL; M13598; AAA60277.1; JOINED.
 DR EMBL; M13880; AAA36580.1; ALT_TERM.
 DR PIR; A25223; TVHURS.
 DR PIR; A24421; TVHURT.
 DR HSP; P08631; IADS.
 DR Genew; HGNC:10261; ROS1.
 DR MIM; 165020; -;
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR003961; FN_III.
 DR InterPro; IPR000033; Ldl_receptor_rep.
 DR InterPro; IPR002011; RTK_kinaseII.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00041; fn3; 7.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00060; FN3; 5.
 DR SMART; SM00135; LY; 2.
 DR SMART; SM00219; TyrcK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
 DR PROSITE; PS00239; RECEPTOR TYR_KIN_II; 1.
 DR PROSITE; PS50011; PROTEIN KINASE_DOM; 1.
 KW Transferase; Tyrosine-protein kinase; Receptor; Transmembrane;
 KW Glycoprotein; ATP-binding; Phosphorylation; Proto-oncogene;
 KW Signal.
 FT SIGNAL 1 27
 FT CHAIN 28 2347
 FT PROTO-ONCOGENE TYROSINE-PROTEIN KINASE
 FT ROS.
 FT EXTRACELLULAR (POTENTIAL).
 FT POTENTIAL.
 FT TRANSMEM 1860 1882
 FT DOMAIN 1883 2347
 FT DOMAIN 1883 2347
 FT DOMAIN 1945 2222
 FT NP_BIND 1951 1959
 FT BINDING 1980 1980
 FT BINDING 1980 1980
 FT MOD_RES 2114 2114
 FT CARBOHYD 52 52
 FT CARBOHYD 114 114
 FT CARBOHYD 123 123
 FT CARBOHYD 324 324
 FT CARBOHYD 352 352
 FT CARBOHYD 396 396
 FT CARBOHYD 471 471
 FT CARBOHYD 607 607
 FT CARBOHYD 628 628
 FT CARBOHYD 706 706
 FT CARBOHYD 714 714
 FT CARBOHYD 732 732
 FT CARBOHYD 939 939
 FT CARBOHYD 961 961
 FT CARBOHYD 1015 1015
 FT CARBOHYD 1087 1087
 FT CARBOHYD 1090 1090
 FT CARBOHYD 1095 1095
 FT CARBOHYD 1211 1211
 FT CARBOHYD 1272 1272

FT CARBOHYD 1330 1330 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1458 1458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1461 1461 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1474 1474 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1565 1565 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1669 1669 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1715 1715 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1738 1738 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1808 1808 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 2213 2213 N-> D (IN REF. 2 AND 3).
 FT CONFLICT 2228 2228 QC -> KS (IN REF. 2 AND 3).
 FT CONFLICT 2246 2246 EDGVDICLSDDIM -> KFDSEFFSFRCTVN (IN
 FT REF. 2).
 SQ SEQUENCE 2347 AA; 263956 MW; E14F3DFD410C1D2A CRC64;
 Query Match 12.08; Score 363; DB 1; Length 2347;
 Best Local Similarity 27.5%; Pred. No. 1.5e-13;
 Matches 126; Conservative 73; Mismatches 196; Indels 64; Gaps 20;
 QY 4 ASAASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKD--- 59
 DB 1913 AAGVGLANACYAIHTLPTQBEIENLPAPPREKLTIRLLLSGAFGEVYEG--TAVDILGV 1970
 QY 60 -----VAIKQIE---SESERKAFIVELRQLSRVHNPNVLYGACL--NPVCLVMEYAE 108
 DB 1971 GSGEIKVAVKTKKSGTDQEKIEFLKEAHLMSKFNHPNLIKQLGVCLLNEPQYIILELME 2030
 QY 109 GGSLYNVLHGAEPYPIY----TAAHMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLL 164
 DB 2031 GGDLLTYLRKARMATFYGPLLTLDVLDLCYDISKGCYVLERMH---FIHRDLAARNCLV 2087
 QY 165 ---VAGGTVLKICDFGTACDI-OTHTMTNNGS---AAMPAEVPFGSGSYSEKCVFSW 215
 DB 2088 SVKDYTSPIVKIGDFGLARDIYKNDYRKRGEGLLPVRWAPESLMDGIFTTQSDVWSP 2147
 QY 216 GIIWLVIT-RRKPFDEIGGPA---FRIMWAVHNGTR--PPLIKNLKPKIESLMTRCWSK 269
 DB 2148 GILIWEILTLGHQPY-----PAHSNLDVLNVYQTGGRLEPP--RNCDDLLWNLMTQCWAQ 2200
 QY 270 DPSQPSMEEIVKIMTHLMRYFPQGADEPLQYPCQYSDEGQSNATSIGSFM----DIAS 325
 DB 2201 EPDQRPFTFHRIQNLQQLFRNFFLNS-----IYQCR--DEA--NNSGVINESFEGDGDV 2253
 QY 326 NTSNKSDDTNMEQVPATNDTIKRLSKLLKNOAKQSQESGRLSLGASHGSSVSELPPTSEG 385
 DB 2254 N-----SDDIMPVVLVLMETKNRGLNVMVLATCGGQGEKSEGLPGSQES---ESCGLRKEE 2306
 QY 386 KRMSADMSEIEARIAATTNGQPPRRRSIQDLTVTGTGTEFG 424
 DB 2307 KEPHADKDFCQEKQVAYCPSGKPEGLNYACLTHSGYGDG 2345

Search completed: December 9, 2002, 22:48:21
 Job time : 32.1274 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:31:39 ; Search time 60.9474 Seconds
(without alignments)
1265.881 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKQLEVRISQQKQKQTS 579

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: A_Geneseq_101002.*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	3014	100.0	579	18 AAW27093	Human transforming
2	3014	100.0	579	20 AAY09542	Human TAK1 protein
3	3014	100.0	579	21 AAY91000	Human TAK-1 protein
4	3014	100.0	590	20 AAY09547	Human TAK1-6XHis p
5	3006	99.7	579	20 AAY28996	Human TGF-beta act
6	3006	99.7	579	23 ABB85033	Pain regulated pro
7	2986	99.1	579	18 AAW27092	Mouse transforming
8	2982.5	99.0	606	20 AAY28997	Human TGF-beta act
9	2746	91.1	567	20 AAY28998	Human TGF-beta act
10	1310	43.5	261	22 AAU33191	Novel human secret

11	966	32.1	678	22	ABB58061	Drosophila melanog
12	505	16.8	252	22	ABB60985	Drosophila melanog
13	482	16.0	473	22	AAW25322	Human protein sequ
14	481	16.0	455	21	AAW18657	A human regulator
15	481	16.0	455	21	AAW83278	Human survival reg
16	481	16.0	455	21	AAW84321	A human cardiovasc
17	462	15.3	1021	23	ABF61000	Novel human protei
18	461.5	15.3	719	22	AAW85513	Human protein kina
19	461.5	15.3	1036	23	ABB80923	Novel human protei
20	459.5	15.2	847	23	AAW22763	Human mitogen acti
21	456.5	15.1	800	22	AAW71957	Human TGF-beta rec
22	456.5	15.1	800	22	AAW65673	Novel protein kina
23	451.5	15.0	1097	23	AAW21717	Human PKIN-12 prot
24	451	15.0	1046	22	AAE11775	Human kinase (PKIN
25	436.5	14.5	1020	22	ABB58999	Drosophila melanog
26	434.5	14.4	367	21	AAW32053	Arabidopsis thalia
27	434.5	14.4	369	21	AAW32052	Arabidopsis thalia
28	434.5	14.4	407	21	AAW32051	Arabidopsis thalia
29	434	14.4	349	22	AAW75571	Human colon cancer
30	434	14.4	369	21	AAW22172	Arabidopsis thalia
31	434	14.4	374	21	AAW22171	Arabidopsis thalia
32	434	14.4	412	21	AAW22170	Arabidopsis thalia
33	427.5	14.2	341	21	AAW25600	Arabidopsis thalia
34	427.5	14.2	391	21	AAW25599	Arabidopsis thalia
35	426.5	14.2	589	21	AAW45984	Arabidopsis thalia
36	426.5	14.2	732	21	AAW45983	Arabidopsis thalia
37	426.5	14.2	760	21	AAW45982	Arabidopsis thalia
38	424	14.1	977	22	ABW71694	Drosophila melanog
39	418	13.9	835	21	AAW01470	Human CARK (Cardia
40	418	13.9	835	21	AAW65674	Novel protein kina
41	418	13.9	928	22	ABG16533	Novel human diagno
42	414	13.7	835	21	AAW01474	Rat CARK (Cardiac
43	412	13.7	859	16	AAW82886	Human leucine-zipp
44	412	13.7	859	18	AAW31227	Human leucine-zipp
45	405	13.4	888	23	ABB57049	Mouse ischaemic co

ALIGNMENTS

RESULT 1
AAW27093
ID AAW27093 standard; Protein; 579 AA.
XX
AC AAW27093;
XX
DT 19-NOV-1997 (first entry)
XX
DE Human transforming growth factor-beta activated kinase TAK-1.
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Homo sapiens.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
(CHUS) CHUGAI PHARM CO LTD.
(UENO/) UENO N.
XX
WPI; 1997-380171/35.
XX
N-PSDB; AAT85095.
XX
DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transmission system

XX PS Claim 15; Page 13-15; 20pp; Japanese.

XX CC The present sequence represents human transforming growth factor-beta

CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the

CC TAK-1 protein which is involved in the TGF-beta family signal

CC transduction system. TAK-1, also known as activator of MAPK Kinase

CC (AMK-1), is an enzyme which is activated by TGF-beta and bone

CC morphogenetic protein (BMP) and activates MAPK Kinase by

CC phosphorylation.

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 18; Length 579;

Best Local Similarity 100.0%; Pred. No. 1.5e-241;

Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60

DB 1 MSTASAASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60

QY 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120

DB 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120

QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLAKICDFTGAC 180

DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLAKICDFTGAC 180

QY 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVTRRRKPPDEIGGPAFRIM 240

DB 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVTRRRKPPDEIGGPAFRIM 240

QY 241 WAVANGTRPPLIKNLKPKPIESLMTRCMSKDPQSPMSMEIVKIMTHLMRYFGADEPLQY 300

DB 241 WAVANGTRPPLIKNLKPKPIESLMTRCMSKDPQSPMSMEIVKIMTHLMRYFGADEPLQY 300

QY 301 PCQYSDGQSNATSTGSPMDIANTSTNSKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360

DB 301 PCQYSDGQSNATSTGSPMDIANTSTNSKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360

QY 361 SESGRSLGASHGSSVESLPPTSGKMSADMEIEARIAATTGNGQPRRSIODTLTVTG 420

DB 361 SESGRSLGASHGSSVESLPPTSGKMSADMEIEARIAATTGNGQPRRSIODTLTVTG 420

QY 421 TEBGQVSSRSSSPSVMTTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAVLTLDHQL 480

DB 421 TEBGQVSSRSSSPSVMTTSGPTSEKPTRSHPTPDDSTDTNGSDNSIPMAVLTLDHQL 480

QY 481 QPLAPCPNKSMAVFEQHCMAEQEYMKVQTEIATLLQKQELVAFLLDDEKQDQNTSL 540

DB 481 QPLAPCPNKSMAVFEQHCMAEQEYMKVQTEIATLLQKQELVAFLLDDEKQDQNTSL 540

QY 541 VOEHKKLLDENKSLSTYYQCKQLVIRSQQKRGTS 579

DB 541 VOEHKKLLDENKSLSTYYQCKQLVIRSQQKRGTS 579

PH Key Location/Qualifiers

FT CDS 183..1922

FT /'tag= a

XX MO9921010-A1.

XX **DB 29-APR-1999.**

XX

XX 22-OCT-1998; 98MO-JP04796.

XX

XX 22-OCT-1997; 97JD-0290188.

XX

XX (CHUS) CHUGAI SEIYAKU KK.

XX

XX Ohtomo T, Ono K, Tsuchiya M;

XX

DR WPI; 1999-312645/26.

DR N-PSDB; AAX56279.

PT Screening for TGF- beta inhibitory substances, which are useful as

PT drugs for treatment of diseases relating to its disorder

PS Claim 4; Page 155-157; 195pp; Japanese.

XX

CC A method has been developed for screening for substances which inhibit

CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method

CC comprises: (a) contacting the polypeptide in the presence of a sample;

CC and (b) detecting the amount of bound polypeptide, in which the sample

CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming

CC growth factor (TGF)-beta inhibitory substances can be used in drugs for

CC indications e.g. as TGF-beta signal transduction inhibitors or

CC activators, or extracellular matrix protein production enhancement

CC inhibitors or activators, or cell proliferation prevention inhibitors or

CC activators, or monocyte migration inhibitors or activators, or

CC physiological activity induction inhibitors or activators, or

CC immunosuppression inhibitors or activators, or amyloid beta protein

CC precipitation inhibitors or activators, and such substances can also be

CC inhibitors of the TAK1 polypeptide function, particularly kinase

CC activity. The present sequence represents human TAK1.

XX

SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 20; Length 579;

Best Local Similarity 100.0%; Pred. No. 1.5e-241;

Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60

DB 1 MSTASAASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVRGAFGVVCAKAKRAKDV 60

QY 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120

DB 61 AIKQIESERKAFIVELRQLSRVNHPIVKLYGACLPVCLVMEYAEGLYVNLHGAE 120

QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLAKICDFTGAC 180

DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLAKICDFTGAC 180

QY 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVTRRRKPPDEIGGPAFRIM 240

DB 181 DIOTHTNNKGSAAWMAPEVEEGSNSEKCDVFSWGIILMEVTRRRKPPDEIGGPAFRIM 240

QY 241 WAVANGTRPPLIKNLKPKPIESLMTRCMSKDPQSPMSMEIVKIMTHLMRYFGADEPLQY 300

DB 241 WAVANGTRPPLIKNLKPKPIESLMTRCMSKDPQSPMSMEIVKIMTHLMRYFGADEPLQY 300

QY 301 PCQYSDGQSNATSTGSPMDIANTSTNSKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360

DB 301 PCQYSDGQSNATSTGSPMDIANTSTNSKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360

QY 361 SESGRSLGASHGSSVESLPPTSGKMSADMEIEARIAATTGNGQPRRSIODTLTVTG 420

DB 361 SESGRSLGASHGSSVESLPPTSGKMSADMEIEARIAATTGNGQPRRSIODTLTVTG 420

QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 Db 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 QY 481 OPLAPCNSKESMAVFEQHCCKMAQYMKVQTEIALLQKQELVAELDDQDKQOQNTSRL 540
 Db 481 OPLAPCNSKESMAVFEQHCCKMAQYMKVQTEIALLQKQELVAELDDQDKQOQNTSRL 540
 QY 541 VOEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579
 Db 541 VOEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579

RESULT 3
 AAY91000
 ID AAY91000 standard; Protein; 579 AA.
 AC AAY91000;
 XX
 DT 04-SEP-2000 (first entry)
 XX
 DE Human TAK-1 protein sequence SEQ ID NO:2.
 XX
 KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
 KW screening; signal transduction; inhibition; inflammatory cytokine;
 KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
 KW antiinflammatory; suppression.
 XX
 OS Homo sapiens.
 XX
 FN WO200023610-A1.
 XX
 PD 27-APR-2000.
 XX
 PF 21-OCT-1999; 99WO-JP05817.
 XX
 PR 21-OCT-1998; 98JP-0299962.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Tsuchiya M, Ontomo T, Sugamata Y, Matsumoto K;
 DR WPI; 2000-339707/29.
 DR N-PSDB; AAA39105.
 XX
 PT Method for screening inhibitors of TAK1 signal transduction for
 PT suppression of inflammatory cytokine production and use as
 PT antiinflammatory agents -
 XX
 PS Example 1; Page 80-84; 100pp; Japanese.
 XX
 CC The present invention describes a method for screening compounds for
 CC inhibition of inflammatory cytokine signal transduction by contacting
 CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
 CC of TAK1/TAB1 binding. Also described is a method for screening compounds
 CC for inhibition of inflammatory cytokine signal transduction in which the
 CC inhibition of TAK1 phosphorylation is selected for; and drug
 CC compositions for the treatment of inflammatory disorders containing as
 CC active component an inflammatory cytokine signal transduction inhibitor.
 CC TAK1 is an essential component of the signalling process which results
 CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
 CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
 CC for the selection of effective antiinflammatory agents. The present
 CC sequence represents human TAK-1, which is used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 579 AA;

Query Match 100.0%; Score 3014; DB 21; Length 579;
 Best Local Similarity 100.0%; Pred. No. 1.5e-241;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASAASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
 Db 1 MSTASAASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
 QY 61 AIKQIESSESEKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYABGGSLYNVLHGAE 120
 Db 61 AIKQIESSESEKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYABGGSLYNVLHGAE 120
 QY 121 PLUPYYTAAHAMSWCICQSQGVAYLHSMOPKALIHRLDKPPNLLLVAGTGLVKICDFGTAC 180
 Db 121 PLUPYYTAAHAMSWCICQSQGVAYLHSMOPKALIHRLDKPPNLLLVAGTGLVKICDFGTAC 180
 QY 181 DIQTHMTNNKGSAAWMAPEVFEKGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 Db 181 DIQTHMTNNKGSAAWMAPEVFEKGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPKPIESLMTRCNSKDPSPSMEIEIVKIMTHLMRYFPGADEPLOY 300
 Db 241 WAVHNGTRPPLIKNLKPKPIESLMTRCNSKDPSPSMEIEIVKIMTHLMRYFPGADEPLOY 300
 QY 301 PCQYSDGQNSATSGSFMDIASNTNSKSDTNWEOVPATNDTIKRLSKLLKNOAKQQ 360
 Db 301 PCQYSDGQNSATSGSFMDIASNTNSKSDTNWEOVPATNDTIKRLSKLLKNOAKQQ 360
 QY 361 SESGRSLSLGASHGSSVESLPPTSEGRKMSADMSIEARIAAATNGQPRRSIQDLTVTG 420
 Db 361 SESGRSLSLGASHGSSVESLPPTSEGRKMSADMSIEARIAAATNGQPRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 Db 421 TEPQVSSRSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
 QY 481 OPLAPCNSKESMAVFEQHCCKMAQYMKVQTEIALLQKQELVAELDDQDKQOQNTSRL 540
 Db 481 OPLAPCNSKESMAVFEQHCCKMAQYMKVQTEIALLQKQELVAELDDQDKQOQNTSRL 540
 QY 541 VOEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579
 Db 541 VOEHKLLDENKSLSTYYQCKKQLEVIRSQOQKRGTS 579

RESULT 4
 AAY09547
 ID AAY09547 standard; Protein; 590 AA.
 XX
 AC AAY09547;
 XX
 DT 21-JUL-1999 (first entry)
 XX
 DE Human TAK1-6xHis protein.
 XX
 KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.
 XX
 OS Homo sapiens.
 OS Synthetic.
 XX
 PN WO9921010-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 22-OCT-1998; 98WO-JP04796.
 XX
 PR 22-OCT-1997; 97JP-0290188.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Ontomo T, Ono K, Tsuchiya M;
 XX
 DR WPI; 1999-312645/26.
 DR N-PSDB; AAX56285.
 XX
 PT Screening for TGF- beta inhibitory substances, which are useful as

PT drugs for treatment of diseases relating to its disorder
 XX
 PS Example 1; Page 171-174; 195pp; Japanese.
 XX

CC A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents TAK1-6xHis from an example of
 CC the present invention.
 CC
 XX
 XX

SQ Sequence 590 AA;

Query Match 100.0%; Score 3014; DB 20; Length 590;
 Best Local Similarity 100.0%; Pred. No. 1.5e-241;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMTEAPSOVLNFEIDYKEIIEVEEVGARGAVGCAKMRKADY 60
 DB 1 MSTASASSSSSSSAGEMTEAPSOVLNFEIDYKEIIEVEEVGARGAVGCAKMRKADY 60
 QY 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGA 120
 DB 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGA 120
 QY 121 PLPYTYAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKCDFGTAC 180
 DB 121 PLPYTYAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKCDFGTAC 180
 QY 181 DIQHTMTNNGKSAAMMAPEVEFGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNGKSAAMMAPEVEFGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLPKPIBSLMTRCWSDQSPMSMEIIVKIMTHLMRYFPGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLPKPIBSLMTRCWSDQSPMSMEIIVKIMTHLMRYFPGADEPLQY 300
 QY 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLKNQAKQ 360
 DB 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLKNQAKQ 360
 QY 361 SEGGRSLGSHGSSVESLPTSEGKRMASDMEIEARIAATGNGQPRRSIQDITLVG 420
 DB 361 SEGGRSLGSHGSSVESLPTSEGKRMASDMEIEARIAATGNGQPRRSIQDITLVG 420
 QY 421 TEBGOVSSRSRSPSVNRITTSPTSEKPTRSHPTPDSDTNGSDNSIPMAYLTLDHQL 480
 DB 421 TEBGOVSSRSRSPSVNRITTSPTSEKPTRSHPTPDSDTNGSDNSIPMAYLTLDHQL 480
 QY 481 QPLAPCNSKESMAVFEQHKMAQEVYKQVTEITALLQKQELVALLDQEKQNTSRL 540
 DB 481 QPLAPCNSKESMAVFEQHKMAQEVYKQVTEITALLQKQELVALLDQEKQNTSRL 540
 QY 541 VOEHKKLLDENKSLSTYVQCKKQLEVIKSGQOQROOTS 579
 DB 541 VOEHKKLLDENKSLSTYVQCKKQLEVIKSGQOQROOTS 579

RESULT 5
 AAY28996
 ID AAY28996 standard; Protein; 579 AA.
 XX
 AC AAY28996;

XX 29-OCT-1999 (first entry)
 DT
 XX
 DE Human TGF-beta activated kinase (TAK) 1a amino acid sequence.
 XX

KW Nuclear factor kappa B; NF-KB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1a.
 XX

OS Homo sapiens.

XX
 PN MO3940202-A1.

XX 12-AUG-1999.

XX 02-FEB-1999; 99MO-JP00422.

XX 30-OCT-1998; 98JP-0309316.
 PR 06-FEB-1998; 98JP-0026003.

XX (TANA) TANABE SEIYAKU CO.

PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;

DR WPJ; 1999-494298/41.
 DR N-PSDB; AAX99696.

PT Nuclear factor kappa B activation inhibitors, useful as preventives
 for, e.g. autoimmune diseases

XX Examples; Page 35-39; 49pp; Japanese.

CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the amino acid sequence of human TAK1a
 CC (hTAK1a) protein.
 CC
 XX
 XX

SQ Sequence 579 AA;

Query Match 99.7%; Score 3006; DB 20; Length 579;
 Best Local Similarity 99.8%; Pred. No. 6.8e-241;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMTEAPSOVLNFEIDYKEIIEVEEVGARGAVGCAKMRKADY 60
 DB 1 MSTASASSSSSSSAGEMTEAPSOVLNFEIDYKEIIEVEEVGARGAVGCAKMRKADY 60
 QY 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGA 120
 DB 61 AIKQIESSEKKAFFVELRQLSRVNHPNIVKLYGACLNPCVLMVEYAEAGSLYNVLHGA 120
 QY 121 PLPYTYAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKCDFGTAC 180
 DB 121 PLPYTYAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKCDFGTAC 180
 QY 181 DIQHTMTNNGKSAAMMAPEVEFGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNGKSAAMMAPEVEFGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLPKPIBSLMTRCWSDQSPMSMEIIVKIMTHLMRYFPGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLPKPIBSLMTRCWSDQSPMSMEIIVKIMTHLMRYFPGADEPLQY 300
 QY 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLKNQAKQ 360
 DB 301 PCQYSDEGQNSATSTGSPMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLKNQAKQ 360

QY 361 SESRLSLGASHGSSVESLPTSEKGRMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 DB 361 SESRLSLGASRGSSVESLPTSEKGRMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDGSDNSIPMAYLTLDHQL 480
 QY 481 QPLAPCNSKESMAVFEHQCKMAQVYMKVOTETALLLQKQELVAELDDQKQNTSRL 540
 DB 481 QPLAPCNSKESMAVFEHQCKMAQVYMKVOTETALLLQKQELVAELDDQKQNTSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEIVRSQQKROGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEIVRSQQKROGTS 579
 RESULT 6
 ABB85033
 ID ABB85033 standard; Protein; 579 AA.
 XX AC ABB85033;
 XX DT 16-MAY-2002 (first entry)
 XX Pain regulated protein sequence 28.
 XX Pain; analgesic; gene therapy; neurological disorder;
 KW neurodegenerative disease.
 XX Homo sapiens.
 XX WO200212338-A2.
 XX 14-FEB-2002.
 XX 03-AUG-2001; 2001WO-EP09011.
 XX 03-AUG-2000; 2000DE-1037759.
 XX (CHEF) GRUENENTHAL GMBH.
 XX Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;
 WPI; 2002-257469/30.
 DR N-PSDB; ABL88437.
 XX Identifying pain-regulating compounds, useful for treating chronic pain
 PT and for diagnosis, by measuring binding of compounds to specific
 PT peptides and proteins
 XX Claim 1; Fig 44; 213pp; German.
 XX The invention relates to identifying pain-regulating substances (A)
 CC comprises (i) incubating a test substance with a cell (or preparation
 CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
 CC either binding of the test substance to (B) or some functional parameter
 CC that is altered by this binding. The method is useful for identifying
 CC pain-regulating substances (A) with analgesic activity. (A) along with
 CC nucleic acid (ABL88411-ABL88441) that encode proteins (B)
 CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
 CC nucleic acid; antibodies against (B); cells that express (B) and agents
 CC that bind to (B), are all useful for treating pain, particularly chronic
 CC pain, including use in gene therapy. The same materials can also be used
 CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
 CC present sequence is that of a polypeptide of the invention.
 XX Sequence 579 AA;
 SQ
 Query Match 99.7%; Score 3006; DB 23; Length 579;
 Best Local Similarity 99.8%; Pred. No. 6.8e-241;
 Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 METASAAASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCAKWKRAKV 60
 DB 1 METASAAASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCAKWKRAKV 60
 QY 61 AIKQIESSESERKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYAEAGGSLYNNVLHGA 120
 DB 61 AIKQIESSESERKAFIVELRQLSRVNHPIVKLYGACLNPCVLMVEYAEAGGSLYNNVLHGA 120
 QY 121 PLPYYTAAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKI CDFGTAC 180
 DB 121 PLPYYTAAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKI CDFGTAC 180
 QY 181 DIQHTMTNNKGSAAWMADEVFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 DB 181 DIQHTMTNNKGSAAWMADEVFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
 QY 241 WAVHNGTRPPLIKNLKPPIESLMTTRCWSKQSPSMEIIVKIMTHLMRYFPFGADEPLQY 300
 DB 241 WAVHNGTRPPLIKNLKPPIESLMTTRCWSKQSPSMEIIVKIMTHLMRYFPFGADEPLQY 300
 QY 301 PCQXSDEGQSNASATSGSFMDIASNTNSKSDTNNEQVPATNDTIKLESKLLKNQAKQQ 360
 DB 301 PCQXSDEGQSNASATSGSFMDIASNTNSKSDTNNEQVPATNDTIKLESKLLKNQAKQQ 360
 QY 361 SESRLSLGASHGSSVESLPTSEKGRMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 DB 361 SESRLSLGASHGSSVESLPTSEKGRMSADMSIEARIAATTGNGOPRRRSIQDLTVTG 420
 QY 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDGSDNSIPMAYLTLDHQL 480
 DB 421 TEPQVSSRSSPSVRMITTSPTSEKPTRSHPWTPDDSTDGSDNSIPMAYLTLDHQL 480
 QY 481 QPLAPCNSKESMAVFEHQCKMAQVYMKVOTEIALLLQKQELVAELDDQKQNTSRL 540
 DB 481 QPLAPCNSKESMAVFEHQCKMAQVYMKVOTEIALLLQKQELVAELDDQKQNTSRL 540
 QY 541 VQEHKLLDENKSLSTYYQCKKQLEIVRSQQKROGTS 579
 DB 541 VQEHKLLDENKSLSTYYQCKKQLEIVRSQQKROGTS 579
 RESULT 7
 AAW27092
 ID AAW27092 standard; Protein; 579 AA.
 XX AC AAW27092;
 XX DT 19-NOV-1997 (first entry)
 XX Mouse transforming growth factor-beta activated kinase TAK-1.
 DE TGF-beta; signal transduction; TGF-beta activated kinase;
 KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
 KW protein kinase.
 XX Mus musculus.
 OS JP09163990-A.
 XX 24-JUN-1997.
 XX 27-SEP-1996; 96JP-0256747.
 XX 29-JUL-1996; 96US-0685625.
 PR 24-SEP-1995; 95JP-0253549.
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA (UENO/) UENO N.
 XX WPI; 1997-380171/35.
 DR N-PSDB; AAT85094.
 XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1

PT - useful for studying the TGF-beta signal transmission system
XX
PS Claim 14; Page 10-12; 20pp; Japanese.
XX
CC The present sequence represents mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transmission system. TAK-1, also known as activator of MAPK kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
XX
XX
SQ Sequence 579 AA;
Query Match 99.1%; Score 2986; DB 18; Length 579;
Best Local Similarity 99.1%; Pred. No. 3,1e-239;
Matches 574; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
DB 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
QY 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACINPVCLVMEYABGSLYNVLHGAE 120
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACINPVCLVMEYABGSLYNVLHGAE 120
DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFGTAC 180
QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFGTAC 180
QY 181 DIQHTMTNKKSAAMWAPVEFGSNYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNKKSAAMWAPVEFGSNYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPRMEIIVKIMTHLMKYPFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPRMEIIVKIMTHLMKYPFGADEPLQY 300
QY 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
DB 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPTPTSEGRKMSADMESEIARIVATAGNQPPRRRIODLTVTG 420
DB 361 SESGRSLGASHGSSVESLPTPTSEGRKMSADMESEIARIVATAGNQPPRRRIODLTVTG 420
QY 421 TEGQVSSSRSSPSVMTTSGPTSEKPRSHWPTDSDTNGSDNSIPMAYLTLDHQL 480
DB 421 TEGQVSSSRSSPSVMTTSGPTSEKPRSHWPTDSDTNGSDNSIPMAYLTLDHQL 480
QY 481 QPLAPCPNSKESMAVEFQHKMAQEVYKQTEITALLQKQELVALDDEKDOQNTSRL 540
DB 481 QPLAPCPNSKESMAVEFQHKMAQEVYKQTEITALLQKQELVALDDEKDOQNTSRL 540
QY 541 VOEHKKLLDENKSLSTYYQCKKQLEVINSOQKRGTS 579
DB 541 VOEHKKLLDENKSLSTYYQCKKQLEVINSOQKRGTS 579

RESULT 8
AAAY28997
ID AAAY28997 standard; Protein; 606 AA.
XX
XX AAAY28997;
XX
DT 29-OCT-1999 (first entry)
XX
XX
DE Human TGF-beta activated kinase (TAK) 1b amino acid sequence.
XX
KW Nuclear factor kappa B; NF-KB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1b.

XX
OS Homo sapiens.
XX
PN WO9940202-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-JP00422.
XX
PR 30-OCT-1998; 98JP-0309316.
PR 06-FEB-1998; 98JP-0026003.
XX
PA (TANA) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
PI WPI; 1999-494298/41.
XX
DR N-PSDB; AAX99697.
XX
PT Nuclear factor kappa B activation inhibitors, useful as preventives
for, e.g., autoimmune diseases
XX
PS Examples; Page 39-43; 49pp; Japanese.
XX
CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1b
CC (hTAK1b) protein.
XX
SQ Sequence 606 AA;
Query Match 99.0%; Score 2982.5; DB 20; Length 606;
Best Local Similarity 95.4%; Pred. No. 6,6e-239;
Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;
DB 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
QY 1 MSTASASSSSSSAGEMIAPSOVLNFEEDYKEIEVEEVGGAFGVYCKAKMRADY 60
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACINPVCLVMEYABGSLYNVLHGAE 120
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACINPVCLVMEYABGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACINPVCLVMEYABGSLYNVLHGAE 120
QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFGTAC 180
QY 181 DIQHTMTNKKSAAMWAPVEFGSNYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
DB 181 DIQHTMTNKKSAAMWAPVEFGSNYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPRMEIIVKIMTHLMKYPFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPRMEIIVKIMTHLMKYPFGADEPLQY 300
QY 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
DB 301 PCOYSDGQNSATSTGSMFMDIASTNTSNKSDTNMEQVPATNDTIKLESKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPTPTSEGRKMSADMESEIARIVATAGNQPPRRRIODLTVTG 420
DB 361 SESGRSLGASHGSSVESLPTPTSEGRKMSADMESEIARIVATAGNQPPRRRIODLTVTG 420
QY 404 -----GNGQPRRRSTODLTVTGTEGQVSSSRSSPSVMTTSGPTSEKPTSHP 453
DB 421 ILDPVEIIVSNGQPPRRRIODLTVTGTEGQVSSSRSSPSVMTTSGPTSEKPTSHP 480
QY 454 WTPDSDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVEFQHKMAQEVYKQTEI 513

```
Db 481 WTPDDSTDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEQCKMAQYMKVQTEI 540
QY 514 ALLLQKQELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 573
Db 541 ALLLQKQELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 600
QY 574 KROGTS 579
Db 601 KROGTS 606

RESULT 9
AAV28998
ID AAY28998 standard; Protein; 567 AA.
AC AAY28998;
XX
XX 29-OCT-1999 (first entry)
DE Human TGF-beta activated kinase (TAK) 1c amino acid sequence.
XX
XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c.
XX
XX Homo sapiens.
OS
XX WO9940202-A1.
PN
XX 12-AUG-1999.
PD
XX 02-FEB-1999; 99WO-JF00422.
PF
XX 30-OCT-1998; 98JP-0309316.
PR
XX 06-FEB-1998; 98JP-0026003.
PS
XX (TANA ) TANABE SEIVAKU CO.
PA
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
PI
XX WPI; 1999-494298/41.
DR
XX N-PSDB; AAX99698.
DR
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases
XX
XX Examples; Page 43-46; 49pp; Japanese.
PS
XX The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1c
CC (hTAK1c) protein.
XX
XX Sequence 567 AA;
SQ

Query Match 91.1%; Score 2746; DB 20; Length 567;
Best Local Similarity 88.8%; Pred. No. 2.7e-219;
Matches 538; Conservative 0; Mismatches 2; Indels 66; Gaps 2;

QY 1 MSTASAASSSSSSAGMIEAPQVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
Db 1 MSTASAASSSSSSAGMIEAPQVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120
Db 61 AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120
```

```
QY 121 PLPYTTAAHAMSWCLQCSQGQVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDFTAC 180
Db 121 PLPYTTAAHAMSWCLQCSQGQVAYLHSMQPKALIHRLDKPPNLLVAGTGLKICDFTAC 180
QY 181 DIOTHTNNGKSAAMAPPEVPEGSGNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
Db 181 DIOTHTNNGKSAAMAPPEVPEGSGNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKQSPRSMEIEIVKIMTHLMRYFPFGADEPLQY 300
QY 301 PCQYSDGQSN SATSGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQQ 360
Db 301 PCQYSDGQSN SATSGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLESKLLKNQAKQQ 360
QY 361 SESGRSLGASHGSSVESLPTTSEGKMSADMSIEARIAATT----- 403
Db 361 SESGRSLGASHGSSVESLPTTSEGKMSADMSIEARIAATTAYSKPKRGRHRTASFGN 420
QY 404 -----GNGOPRRRSIQDLTVTGTGPGOVSSRSSPSVRMITTSGPTSEKPTRSHP 453
Db 421 ILDPVEIVISGNGOPRRRSIQDLTVTGTGPGOVSSRSSPSVRMITTSGPTSEKPTRSHP 480
QY 454 WTPDDSTDTNGSDNSIPMAYLTLDHQLQPLAPCPNSKESMAVFEQCKMAQYMKVQTEI 513
Db 481 WTPDDSTDTNGSDNSIPMAYLTLDHQLQ----- 508
QY 514 ALLLQKQELVAELDQDEKQDQNTSRLVQEHKLLDENKSLSTYYQCKKQLEVIRSQQQ 573
Db 509 -----QELVAELDQDEKQDQNTSRLVQEHKLLDENKGLSTYYQCKKQLEVIRSQQQ 561
QY 574 KROGTS 579
Db 562 KROGTS 567

RESULT 10
AAU33191
ID AAU33191 standard; Protein; 261 AA.
XX
XX AC AAU33191;
XX
DT 18-DEC-2001 (first entry)
XX
DE Novel human secreted protein #3682.
XX
XX Human; vaccination; gene therapy; nutritional supplement;
KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
OS Homo sapiens.
XX
XX WO200179449-A2.
PN
XX 25-OCT-2001.
PD
XX 16-APR-2001; 2001WO-US08656.
PF
XX 18-APR-2000; 2000US-0552929.
PR 26-JAN-2001; 2001US-0770160.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-611725/70.
XX
XX Nucleic acids encoding a range of human polypeptides, useful in genetic
PT vaccination, testing and therapy -
XX
XX Claim 20; Page 726; 765pp; English.
```

XX The invention relates to novel human secreted polypeptides. The
CC polypeptides and antibodies to the polypeptides are useful for
CC determining the presence of or predisposition to a disease associated
CC with altered levels of polypeptide. The polypeptides are also useful for
CC identifying agents (agonists and antagonists) that bind to them. Cells
CC expressing the proteins are useful for identifying a therapeutic agent
CC for use in treatment of a pathology related to aberrant expression or
CC physiological interactions of the polypeptide. Vectors comprising
CC the nucleic acids encoding the polypeptides and cells genetically
CC engineered to express them are also useful for producing the proteins.
CC The proteins are useful in genetic vaccination, testing and
CC therapy, and can be used as nutritional supplements. They may be used to
CC increase stem cell proliferation; to regulate haematopoiesis; and in
CC bone, cartilage, tendon and/or nerve tissue growth or regeneration;
CC immune suppression and/or stimulation; as anti-inflammatory agents; and
CC in treatment of leukemias. AAU29510-AAU3304 represent the amino acid
CC sequences of novel human secreted proteins of the invention.

XX Sequence 261 AA;

Query Match 43.5%; Score 1310; DB 22; Length 261;

Best Local Similarity 99.6%; Pred. No. 1.8e-100; Mismatches 1; Indels 0; Gaps 0;

Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 MDIASTNTSNKSDPTNMEQVPAINDTIKRLSKLLKNQAQOSEGSLSGASHGSSVESL 60

QY 380 PPTSEGRKMSADMSIEIARIATTTNGQPRRSIIDLTVGTGEPGVSSRSSPSVRMT 439

DB 61 PPTSEGRKMSADMSIEIARIATTTNGQPRRSIIDLTVGTGEPGVSSRSSPSVRMT 120

QY 440 TSPPTSEKPRSHPMPTPDDSTDTNGSDNSIPMAVYLTLDHQLQPLACPSKESMAVFEQH 499

DB 121 TSPPTSEKPRSHPMPTPDDSTDTNGSDNSIPMAVYLTLDHQLQPLACPSKESMAVFEQH 180

QY 500 CKMAOEYKQVTEIALLQKQELVAELDDKQDQNTSLVQEHKLLDENKSLSTYYQ 559

DB 181 CKMAOEYKQVTEIALLQKQELVAELDDKQDQNTSLVQEHKLLDENKSLSTYYQ 240

QY 560 QCKKOLEVIRSQOQKRGTS 579

DB 241 QCKKOLEVIRSQOQKRGTS 260

RESULT 11

ABBS8061

ID ABBS8061 standard; Protein; 678 AA.

XX ABBS8061;

DT 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 975.

XX Drosophila; developmental biology; cell signalling; insecticide;

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PMD, Myers EW;

XX WPI; 2001-656860/75.
DR N-PSDB; ABL02164.
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure; SEQ ID NO 975; 21pp + Sequence listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABBS7737-ABBS72072).

CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 678 AA;

Query Match 32.1%; Score 966; DB 22; Length 678;

Best Local Similarity 36.0%; Pred. No. 3e-71; Mismatches 207; Indels 98; Gaps 14;

Matches 226; Conservative 97; Mismatches 207; Indels 98; Gaps 14;

QY 18 MIEAPQVLNPEEIDYKEIEVEEYVGRGAFQVCKAKWRAKDAVAIKQIESSEKKAFFIVE 77

DB 1 MATASLDALQAAVDFSEITLREKVGSGYGVCKAAMRKALYAKVEFFASAKQDIKE 60

QY 78 LRQSRVNHENIVKLYG-ACLNPVCLMEYABEGSLYLVHGAEPPLPYTAANMSWCL 135

DB 61 VKQSRVGHENIILHGISYQKATYLLIMEFAGGSLHNFHG-KVXPAYSLHAMSWMAR 119

QY 136 QCSQGVAYLHMQKALIRHDLKRPNLLVAGGVTLKICPFGTRACDIQTMTNKGSAAM 195

DB 120 QCAAGLAVYLAHMTPKPLIHDVFKPLNLLTNKGKNLKICDFGVADASTMTNKGSAAM 179

QY 196 MAPEYFEGSNSEKCDVFSNGIILMEVITRRKPEDEGSAFRIMAVHNGTSPRLIKNL 255

DB 180 MAPEYFEGSKTEKCDIFSMALIVMEVLRSKQPKGIDN-AITTIQMKIYAGERPPLITTC 238

QY 256 PKPIESLMTRCXKDPSPQSPMEBIVKIMTHLMKRYPPGADEPLYQ----- 300

DB 239 PKRIEDLMTACWKTVPEDRPSMQYIVGMHEIVKDYTGADKALEYTFVNGQIVTKESDGT 298

QY 301 -----PCOYSDGQSN-----SATSTGSPMDIATNTSNKSDTN 334

DB 299 VAAQPDLSLSQEGELSPSSQTLPPTAANNANVNAIAISKTTTSMTEINTSSTDITPTN 358

QY 335 MEQVP-----ATN--DTIKRLSKLLKN-----QAKQOSEGRLSLGASHSSV 376

DB 359 SGQLDNNPDLFMYVTNRWDALPEESNESRNDSEFNLTSSAATQKLELTRNGMLMACKPM 418

QY 377 ESLPPTSEKMSADMSIEIARIATTTNGQPRRSIIDLTVGTGP-----GQ 425

DB 419 EQL--TLIDVEANGFDLPSSESSSSSTNAKSDGR-----LTVIDTKPVMTTDLNNGG 472

QY 426 VSSRSSPSVRMTTSGPTSEKPTRSHPW-----TPDDSTDTNGSDNSIPMAVLT 476

DB 473 IHAHNGILSHANGWQARDELQOEHEQELIVSLDVDPDDEENDGTQSLAE--IL 529

QY 477 DHOLOPLAPCNPSKESMAVFEQHCKMAOEYKQVTEIALLQKQELVAELDDKQDQCN 536

DB 530 DPLOPEPPIPNDAESQIYRDRHRMAKEYLSVDTNLVYADFFDKLIVMDRTERQK- 588

QY 537 TSLRVOEHKLLDENKSLSTYYQCKKQ 564

DB 589 -----QELLRKMKDKEGQLYNNLQOQ 611

```

RESULT 12
ABB60985
ID ABB60985 standard; Protein; 252 AA.
XX
AC ABB60985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEXE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
DR N-PSDB; ABL05088.
XX
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT Genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 9747; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC (ABB57737-ABB72072).
CC
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 252 AA;

Query Match 16.8%; Score 505; DB 22; Length 252;
Best Local Similarity 40.6%; Pred. No. 1.4e-33;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

QY 29 EIDYKEIEVEYVGARGVWCCKAKRQDVAKIQIESERKAPFIVELRLQSLRVNHPN 88
Db 6 EGVPEIEIOTKELIGTFYGVSVYRAVRNRREITALKIRBGCEKKTIERIYQITKASHVN 65
QY 89 IVKLYGACLNPCV--LVMYAEAGGSLYNVLHGAEPLPYTTAAHAMSWCLQCSGVAYLHS 146
Db 66 IVELYGTSRHEGCALLLMEFVDGGSLSSEFLH-AKSPSYSHAHAFNWAHQIAGGIAYLHG 124
QY 147 MQPKALIHRLDKPNNLLVAGGTVLKI DFGTACD:OTHTMNTNKGSAWMAPE----- 199
Db 125 MQPKAVIHRDKLPNTLLECKGLKLIKIDFGTVVDUSQISNAGTCRYKAEVRELPDF 184
QY 200 -----VPEGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRINWAVH 244
Db 185 KSNRIINQPTGFKVLQGNKPKCDVYSWAIITFWELLSRKEPPEQY-NTLIFELYMAIN 243
QY 245 NGTR 248
|

```

```

Db 244 EGKR 247

RESULT 13
AAM25322
ID AAM25322 standard; Protein; 473 AA.
XX
AC AAM25322;
XX
DT 16-OCT-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:837.
XX
KW Human; cancer; ulcer; HIV infection; human immunodeficiency virus;
KW antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;
KW antibacterial; endocrine; cardiant; central nervous system; virucide;
KW anti-HIV; fungicide; antimutagen; cardiovascular; antianaemic; anaemia;
KW antiaggregant; haemostatic; vulnery; antiulcer; osteopathic; eczema;
KW dermatological; anti allergic; antiasthmatic; antidiabetic; cytostatic;
KW neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;
KW immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;
KW antianaphylactic; rheumatoid arthritis; septic shock; pancreatitis;
KW cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;
KW genetic disease; haematopoietic disorder; platelet disorder; asthma;
KW thrombocytopaenia; osteoporosis; severe combined immunodeficiency;
KW allergic rhinitis; diabetes; multiple sclerosis; depression;
KW Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;
KW neurological disorder.
XX
XX Homo sapiens.
OS
XX
PN WO200153455-A2.
XX
PD 26-JUL-2001.
XX
PF 22-DEC-2000; 2000WO-US35017.
XX
PR 23-DEC-1999; 99US-0471275.
PR 21-JAN-2000; 2000US-0488725.
PR 25-APR-2000; 2000US-0552317.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
PI Tang YT, Liu C, Drmanac RT;
XX
WPI; 2001-457603/49.
DR N-PSDB; AAM99263.
XX
XX
PT Isolated human polynucleotides encoding polypeptides, useful for the
PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection -
XX
XX Claim 20; Page 191; 1217pp; English.
XX
CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to
CC AAM25963. The proteins can have activities based on the tissues and
CC cells they are expressed in, such as: antinflammatory; antirheumatic;
CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant;
CC central nervous system; virucide; anti-HIV; fungicide; antimutagen;
CC cardiovascular; antianaemic; antiaggregant; haemostatic; vulnery;
CC antiulcer; osteopathic; dermatological; anti allergic; antiasthmatic;
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides
CC encoding them can be used in gene therapy, antisense therapy and vaccine
CC production. The proteins and polynucleotides are useful for screening for
CC agonists or antagonists of a protein and for the treatment and diagnosis
CC of disorders associated with the activity of a protein e.g. inflammation,
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,
CC osteoporosis, severe combined immunodeficiency, eczema, allergic
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and

```


CC other developmental disorders of the central nervous system,
CC neuromuscular disorders, myasthenia gravis, cell proliferative disorders
CC such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
CC including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
CC bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
CC inflammatory disorder such as Addison's disease, acquired
CC immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
CC rheumatoid arthritis, microbial infection and trauma.
XX
SQ Sequence 455 AA;

Query Match 16.0%; Score 481; DB 21; Length 455;
Best Local Similarity 30.4%; Pred. No. 3.3e-31;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
QY 27 NFEEDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDLQFFENGCGGSGSVYRAKWISQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHENIVKLYGACINP--VCLVMEYAEGLSYLVNHLG--AEPLPYTAAHMSWCILQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMTWATDV 114
QY 138 SOGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKIICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVTHRLDKSRNVVIAADG-VLKICDPG-ASRFNHTTHMSLVGTTPW 172
QY 196 MAPEVFEKSYSEKCDVFSWGIILWEVITRRKPPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSGVVLWEMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230
QY 255 LKPIESLMTRCWSKDPORSOPGMEETVKMTLHMRYPFGADEPLOVPCQYSDGOSNSAT 314
DB 231 CPRSPFELLHQWEADAKRPFKQIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGLSL---GAS 371
DB 268 ---SLPDKCNSFLHNKAERCE-IEATLERLKLRLDLSPKEQELKERERLKNWEOKLT 323
QY 372 HGSSVESLPPTS-----EGKRMASDMSEIARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EOSNTPLLLPLAARMSBESYFESKTESNSAEMSCQITATSNGBGHGMNPSLQAWMLMGF 383
QY 422 EPGQVSRSSSPSV 435
DB 384 --GDIFSMNKAGAV 395

RESULT 15
AAY83278
ID AAY83278 standard; Protein; 455 AA.
XX
AC AAY83278;
XX
DT 16-AUG-2000 (first entry)
XX
DE Human survival regulating kinase (SRK).
XX
KW Survival regulating kinase; SRK; Raf; MBP; BAD; MAPKK; MAPKK;
KW HAX-1; protein kinase; autophosphorylation; cell growth; regulation;
KW apoptosis; cell survival; nuclear targeting; tumour; human;
KW autoimmune disease.
XX
OS Homo sapiens.
XX
PN WO200022142-A2.
XX
PD 20-APR-2000.
XX
PF 20-SEP-1999; 99WO-US22008.
XX
PR 13-OCT-1998; 98US-0104088.
XX

PA (ONYX-) ONYX PHARM INC.
XX
PI Ruggieri R, Callow M, Diaz P;
XX
DR WPI; 2000-317994/27.
DR N-PSDB; AA293783.
XX
PT Novel human survival regulating kinase polypeptide for screening agents
PT which modulate biological pathways associated with SRK useful in
PT treating autoimmune diseases, tumors and apoptosis-related disorders
XX
XX Claim 4; Figure 2; 62pp; English.
XX
PS Survival regulating kinases (SRK) are a class of proteins involved in
CC cell signal transduction pathways such as mitogen-activated protein
CC kinase pathways. A protein kinase activity means that the SRK can
CC catalyse a reaction in which a phosphate group is transferred from a
CC phosphate donor to a phosphate acceptor amino acid residue,
CC preferably the hydroxyl side chain of a serine or threonine.
CC Substrates for SRK include SRK, MBP and BAD and SRK's protein kinase
CC activity is similar to that of a MAPKK such as Raf. has a range of
CC other activities including a cell growth-regulatory activity, a cell
CC survival promoting activity, a HAX-1 binding activity, an apoptosis
CC suppressing activity a MAPKK activation or stimulatory activity, a
CC nuclear targeting activity and a SRK-specific immunogenic activity.
CC SRK is useful for identifying agents which modulate cellular
CC transformations mediated by Ras and SRK and agents that modulate the
CC apoptosis suppression activity of SRK. This information may be useful
CC in the treatment of autoimmune diseases, tumours and apoptosis
CC related disorders.
XX
SQ Sequence 455 AA;

Query Match 16.0%; Score 481; DB 21; Length 455;
Best Local Similarity 30.4%; Pred. No. 3.3e-31;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
QY 27 NFEEDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDLQFFENGCGGSGSVYRAKWISQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHENIVKLYGACINP--VCLVMEYAEGLSYLVNHLG--AEPLPYTAAHMSWCILQC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLSDYINSRSEEM---DMDHMTWATDV 114
QY 138 SOGVAYLHSMQPKALIHRLDKPNNLLVAGTVLKIICDFTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVTHRLDKSRNVVIAADG-VLKICDPG-ASRFNHTTHMSLVGTTPW 172
QY 196 MAPEVFEKSYSEKCDVFSWGIILWEVITRRKPPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSGVVLWEMLTREVPFKGLEG--LQVAVLVEKNERLTIPSS 230
QY 255 LKPIESLMTRCWSKDPORSOPGMEETVKMTLHMRYPFGADEPLOVPCQYSDGOSNSAT 314
DB 231 CPRSPFELLHQWEADAKRPFKQIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGLSL---GAS 371
DB 268 ---SLPDKCNSFLHNKAERCE-IEATLERLKLRLDLSPKEQELKERERLKNWEOKLT 323
QY 372 HGSSVESLPPTS-----EGKRMASDMSEIARIAATT-GNGQPRRRSIQDLTVTGT 421
DB 324 EOSNTPLLLPLAARMSBESYFESKTESNSAEMSCQITATSNGBGHGMNPSLQAWMLMGF 383
QY 422 EPGQVSRSSSPSV 435
DB 384 --GDIFSMNKAGAV 395

Search completed: December 9, 2002, 22:47:16
Job time : 63.9474 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:47:24 ; Search time 33.1468 Seconds
(without alignments)
1679.251 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAASSSSSSAGEMIE.....OCKKOLEVIRSQOQRQGT 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3006	99.7	579	2 JC5955	transforming growt
2	2982.5	99.0	606	2 JC5956	transforming growt
3	2750	91.2	567	2 JC5957	transforming growt
4	493.5	16.4	954	1 S68178	mixed-lineage prot
5	459.5	15.2	847	1 A53800	probable protein k
6	434.5	14.4	407	2 G84635	protein kinase hom
7	434	14.4	412	2 T10671	protein kinase hom
8	433.5	14.4	1338	2 T18287	protein-tyrosine k
9	427.5	14.2	391	2 T48115	protein kinase ATM
10	421.5	14.0	357	2 C84856	probable protein k
11	418.5	13.9	394	2 JU0229	mixed-lineage prot
12	416	13.8	328	2 T16747	hypothetical prote
13	414	13.7	410	2 B35670	protein-tyrosine k
14	412	13.7	668	2 JC2363	protein kinase (EC
15	412	13.7	1030	2 F96763	hypothetical prote
16	411.5	13.7	546	2 D84555	probable protein k
17	411.5	13.7	553	2 T04683	hypothetical prote
18	408.5	13.6	888	2 A55318	serine/threonine p
19	406.5	13.5	370	2 T46150	protein kinase ATN
20	404.5	13.4	888	2 JC5399	dual leucine zippe
21	398	13.2	821	2 T48400	serine/threonine-p
22	397.5	13.2	829	2 T07406	probable protein k
23	391.5	13.0	736	2 T05137	protein kinase hom
24	391	13.0	364	2 G71410	probable protein k
25	391	13.0	387	2 T22511	hypothetical prote
26	390	12.9	963	2 T09911	probable serine/th
27	389.5	12.9	356	2 S61766	protein kinase ATN
28	387	12.8	475	2 T12955	probable protein k
29	385	12.8	438	2 C86273	protein kinases ho

RESULT 1

JC5955

transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human

C:Species: Homo sapiens (man)

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C:Accession: JC5955

R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.

Biochem. Biophys. Res. Commun. 243, 545-549, 1998

A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind

A:Reference number: JC5955; MUID:98153801; PMID:9480845

A:Accession: JC5955

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-579 <SAK>

A:Cross-references: DDBJ:AB009356; NID:G2924623; PIDN:BA25025.1; PID:G2924624

C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog

C:Keywords: phosphotransferase

Query Match 99.7%; Score 3006; DB 2; Length 579;
Best Local Similarity 99.8%; Pred. No. 1.9e-126;
Matches 578; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASAASSSSSSAGEMIEAPSVLNFEEIDYKEIEVEEVVGRGAFGVCAKWKRAKDV 60
Db 1 MSTASAASSSSSSSSAGEMIEAPSVLNFEEIDYKEIEVEEVVGRGAFGVCAKWKRAKDV 60

QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMVEYAGGSLYNVLHGAE 120
Db 61 AIKQIESESEKAFIVELRQLSRVNHPIVKLYGACLPVCLVMVEYAGGSLYNVLHGAE 120

QY 121 PLPYVYTAHAHMSWCLQCSQGVAYLHSMOPKALIHRLDKPPNLLVAGTVLKI CDGFTAC 180
Db 121 PLPYVYTAHAHMSWCLQCSQGVAYLHSMOPKALIHRLDKPPNLLVAGTVLKI CDGFTAC 180

QY 181 DIQTHMTNKNKSAAMWAPFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQTHMTNKNKSAAMWAPFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTPPLIKNLPKPIESLMTRCWSKDPSPSMEEIVKIMTHLMRYPPGGADEPLQY 300
Db 241 WAVHNGTPPLIKNLPKPIESLMTRCWSKDPSPSMEEIVKIMTHLMRYPPGGADEPLQY 300

QY 301 PCQYSDGQSNATSSTGFMDOIATNTSNKSDTNNEQVPATNDTIKRLSKLLKNQAKQQ 360
Db 301 PCQYSDGQSNATSSTGFMDOIATNTSNKSDTNNEQVPATNDTIKRLSKLLKNQAKQQ 360

QY 361 SESGRLSLGASHGSGVESLPPPTSEGRKMSADMSIEARIAATTNGOPRRRSIODLTVTG 420
Db 361 SESGRLSLGASHGSGVESLPPPTSEGRKMSADMSIEARIAATTNGOPRRRSIODLTVTG 420

QY 421 TEPGVSSRSRSPSVRMITTSGPITSEKPTRSHPTDSDTNGSDNSIPMAYITLHQH 480
Db 421 TEPGVSSRSRSPSVRMITTSGPITSEKPTRSHPTDSDTNGSDNSIPMAYITLHQH 480

ALIGNMENTS

Db 421 TEEGQVSSRSSSVKMITTSGETSEKPTRSHWTPDDSTDTNGSNSIPMAVLTLDHQL 480
QY 481 QPLAPCNSKESNAVEEHCMAQOEYMKVQTEIALLLQKQELVAELDQDEKQNTSL 540
Db 481 QPLAPCNSKESNAVEEHCMAQOEYMKVQTEIALLLQKQELVAELDQDEKQNTSL 540
QY 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQCKRGTS 579
Db 541 VOEHKKLLDENKSLSTYYQCKKQLEVRISQCKRGTS 579

RESULT 2

JC5956
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: JC5956
R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A/Reference number: JC5955; MUID:98153801; PMID:9480845
A/Accession: JC5956
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-606 <SAK>
A/Cross-references: DDBJ:AB009357; NID:92924625; PID:BA25026.1; PID:92924626
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: phosphotransferase

Query Match 99.0%; Score 2982.5; DB 2; Length 606;
Best Local Similarity 95.4%; Pred. No. 2.2e-125;
Matches 578; Conservative 0; Mismatches 1; Indels 27; Gaps 1;
QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVGGAFAVVCAMKRAKDV 60
Db 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVGGAFAVVCAMKRAKDV 60
QY 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVCLMEYAEAGSLYNNVLHGA 120
Db 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVCLMEYAEAGSLYNNVLHGA 120
QY 121 PLPYTTAAHMSWCLQSGVAYLHSMQPAKLIHRDLKPNLLLVAGTVLKI CDGTAC 180
Db 121 PLPYTTAAHMSWCLQSGVAYLHSMQPAKLIHRDLKPNLLLVAGTVLKI CDGTAC 180
QY 181 DIQHTMTNKKSAAMMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQHTMTNKKSAAMMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPRPIESLMTRCWSKQPSQPSMEEIVKIMTHLMRYFPGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLPRPIESLMTRCWSKQPSQPSMEEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQNSATSTGSPMDIATNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIATNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMEIEARIAATTAYSKPKRHKRTASFGN 420
Db 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMEIEARIAATTAYSKPKRHKRTASFGN 420
QY 421 ILDPVEIVISGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVMMITTSGETSEKPTRSH 480
Db 421 ILDPVEIVISGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVMMITTSGETSEKPTRSH 480
QY 481 WTPEDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEEHCMAQOEYMKVQTEI 540
Db 481 WTPEDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEEHCMAQOEYMKVQTEI 540
QY 541 ALLQKQELVAELDQDEKQNTSLVOEHKKLLDENKSLSTYYQCKKQLEVRISQCK 573
Db 541 ALLQKQELVAELDQDEKQNTSLVOEHKKLLDENKSLSTYYQCKKQLEVRISQCK 573

QY 574 KRQGTSS 579
Db 601 KRQGTSS 606

RESULT 3

JC5957
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C/Accession: JC5957
R/Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A/Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A/Reference number: JC5955; MUID:98153801; PMID:9480845
A/Accession: JC5957
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-567 <SAK>
A/Cross-references: DDBJ:AB009358
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: phosphotransferase

Query Match 91.2%; Score 2750; DB 2; Length 567;
Best Local Similarity 88.9%; Pred. No. 4e-115;
Matches 539; Conservative 0; Mismatches 1; Indels 66; Gaps 2;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVGGAFAVVCAMKRAKDV 60
Db 1 MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVGGAFAVVCAMKRAKDV 60
QY 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVCLMEYAEAGSLYNNVLHGA 120
Db 61 AIKQIESSEKKAIVELRQLSRVNHNPNIYKLYGACLNPCVCLMEYAEAGSLYNNVLHGA 120
QY 121 PLPYTTAAHMSWCLQSGVAYLHSMQPAKLIHRDLKPNLLLVAGTVLKI CDGTAC 180
Db 121 PLPYTTAAHMSWCLQSGVAYLHSMQPAKLIHRDLKPNLLLVAGTVLKI CDGTAC 180
QY 181 DIQHTMTNKKSAAMMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
Db 181 DIQHTMTNKKSAAMMAPEVEEGSNSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPRPIESLMTRCWSKQPSQPSMEEIVKIMTHLMRYFPGADEPLQY 300
Db 241 WAVHNGTRPPLIKNLPRPIESLMTRCWSKQPSQPSMEEIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQYSDGQNSATSTGSPMDIATNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
Db 301 PCQYSDGQNSATSTGSPMDIATNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQ 360
QY 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMEIEARIAATTAYSKPKRHKRTASFGN 420
Db 361 SESGRSLIGASHGSSVESLPTSEGRKMSADMEIEARIAATTAYSKPKRHKRTASFGN 420
QY 421 ILDPVEIVISGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVMMITTSGETSEKPTRSH 480
Db 421 ILDPVEIVISGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVMMITTSGETSEKPTRSH 480
QY 481 WTPEDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEEHCMAQOEYMKVQTEI 540
Db 481 WTPEDDSTDTNGSDNSIPMAVLTLDHQLQPLAPCNSKESNAVEEHCMAQOEYMKVQTEI 540
QY 541 ALLQKQELVAELDQDEKQNTSLVOEHKKLLDENKSLSTYYQCKKQLEVRISQCK 573
Db 541 ALLQKQELVAELDQDEKQNTSLVOEHKKLLDENKSLSTYYQCKKQLEVRISQCK 573

RESULT 4

Db	434	-VERELHLLMCQLSQEQKPRVRKRGKFNFKSRLLDKLRGGSHISLPSGFFHKITVQASPTL	492
Qy	383	SEGRKMSADMSSETEARIAATTGNGQRRRSIQDLTVTGTPEPGQVSRSSPSVRMITTSG	442
Db	493	D--KRRKSGDCAASPAPSPSII-----PRLRAIR---LTPVDCGGSSSSGSGSGTWSRGG	542
Qy	443	PTSEK-----PTRSHPWTPDDSS	459
Db	543	PPKKEELVGKKKGRTWGFSST	564
RESULT 5			
A53800			
N;Alternate names: protein kinase PTK1; protein kinase SPRK			
C;Species: Homo sapiens (man)			
C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999			
C;Accession: A53800; I58395			
R;Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.			
J. Biol. Chem. 269, 15092-15100, 1994			
A;Title: Identification and characterization of SPRK, a novel src-homology 3 domain-contain			
A;Reference number: A53800; MUID:94253068; PMID:8195146			
A;Accession: A53800			
A;Status: preliminary			
A;Molecule type: mRNA			
A;Residues: 1-847 <GL>			
A;Cross-references: GB:U07747; NID:9464027; PIDN:AAAL19647.1; PID:9464028			
R;Ing, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassam, N.J.			
Oncogene 9, 1745-1750, 1994			
A;Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domain			
A;Reference number: I58395; MUID:94239754; PMID:8183572			
A;Accession: I58395			
A;Status: preliminary; translated from GB/EMBL/DBJ			
A;Molecule type: mRNA			
A;Residues: 1-847 <RES>			
A;Cross-references: GB:IJ3976; NID:9488295; PIDN:AAA59859.1; PID:9488296			
C;Genetics:			
A;Gene: GDB:MLK3; PTK1; SPRK			
A;Cross-references: GDB:134755; OMIM:600050			
A;Map position: 11q13.1-11q13.3			
C;Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology			
F;48-100/Domain: SH3 homology <SH32>			
F;115-383/Domain: protein kinase homology <KIN>			
F;123-131/Region: protein kinase ATP-binding motif			
F;403-424/Region: leucine zipper motif			
F;438-459/Region: leucine zipper motif			
F;468-482/Region: basic			
Query Match 15.2%; Score 459.5; DB 1; Length 847;			
Best Local Similarity 27.8%; Pred No. 1.7e-13;			
Matches 149; Conservative 88; Mismatches 204; Indels 95; Gaps 19;			
Qy	22	PSQVLNPFEEIDYKEIEVEEVVGARGFVGVCKAKWRADVAIK-----QIESESR	71
Db	108	PCEVASFQ-----ELRLEEVIGIGGFKVYRGSGWRGELVAVRAARDPDEDISVTAESVR	162
Qy	72	KATFVELRQLSRVNHFNIVKLYGACLNLP--VCLVMRYAEGGSLYVNLHGAELPYPYTAH	129
Db	163	Q-----EARLFAMLAHFNITALKAVCLEENPCLCUMVEYAAAGPLSRALAGRRVPVHV-----	214
Qy	130	AMGWCLQCSOGVAYLHSMQPKALIHRLDKPPNLLV-----AGGTVLKICDFGTACD-	181
Db	215	LVNWAVQIARGMHYLHCEALVPVIRHDLKSNILLLOPIESDDMEHKTIXITDFGLAREW	274
Qy	182	IOTHMTNNKSGAAWMAPEVFEGSNYGEKCDVFSWGIIIMEVITRRKPPFDEIGGPAFRIMW	241
Db	275	HKTTQMSAAGTYAWMAPEVIKASTFSKGSVDVMSFGVLLMELLTGTGVPYRGIDCLAVAYGV	334
Qy	242	AVHNGTRTPPLIKNLPRPIESLMTRCWSKDPQSPPS-----MEEIVKIMTHLMRYFPGADEP	297
Db	335	AVNKLTL-PIPTSCPEPFAQLMADMCWAQDPHRRPDPFASIIQQLEALEAQVLEMP-----	388

298 LQYPCQSDGSGNSATSTGSPMDIAS-----TNTSNKSDTMEQVAPATNDT 345
 Db 389 -RDSFHSMOEGAKREI--QGLFDELPAKEXKLLSREELTPRAERQSAEQRLRREHLL 445
 QY 346 KRLSEKLLKNQAK---QSESGRSLGASHGSSVES-LPTSEGKRMASDMSEIARIAA 401
 Db 446 AQWELVEFERELTLLQQVDRRPHVRRRRGTFFKSKLAARDGGERISNPL-DFKRIITV 504
 QY 402 TTGNQCPRRRST-----QDLTVTGTEPGGVSSR-----SSPSVRYMT 439
 Db 505 QASPLDLRRNRNFVEVQPGDSPTPPPFRALQLEPAEGQAMGRQSPRRLESDSNGERRACW 564
 QY 440 TSGPSEKRP-----TRSHPTPPDSDTDNGSNDNISPAAYLTLDHQLOPLAPCPM 488
 Db 565 AMGPSSPKPEGAQNGRRRSR--MDEATWLDSDSSPLGSESTPPALGNPPRPS 617

 RESULT 6
 G84635
 probable protein kinase [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
 C:Accession: G84635
 R:Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffet, K.S.; Cronin, L.A.; Shen, M.; Vancken, S.E.; Umayam, L.; Tallon,
 Eus., D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter,
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; PMID:20083487; PMID:10617197
 A:Accession: G84635
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-407 <STO>
 A:Cross-references: GB:AE002093; NID:94337195; PIDN:AA018109.1; GSPDB:GNO0139
 C:Genetics:
 A:Gene: At2g24360
 A:Map position: 2
 C:Superfamily: kinase-related transforming protein; protein kinase homology

 Query Match 14.4%; Score 434.5; DB 2; Length 407;
 Best Local Similarity 34.8%; Pred. No. 1e-12;
 Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;

 QY 12 SSSAGEMTAPSOV-----LNFEF--IDYKEIVEEVYGRGA 46
 Db 77 SLVSQSVFRPGRVTHALNDALQAQALMDTRYPTGCTVYDEWTIDRLKLNMGPAFAQGA 136
 QY 47 FGVCKAKKARADVAIKQTE--SESERKA-----FIVELPOLSRNHNIVYKLVGACLN 98
 Db 137 FEKLYKGTNGVDVAKILERPBNSEPKQFMEOQFOQEVSMLANIKHNIVRFICACK 196
 QY 99 PV--CLVMEYAEAGSLYNYL---HGAEPLPYTTAAHAMSWCLQCSQGVAYLHSMQPKAL 152
 Db 197 PAVWMCIVTEYAGGSVRQGLPTRQNRNAVPLKL-----AVKQALDVARGMAYVHG---RNF 248
 QY 153 IHRDLKLPALLVAGSTVLKICDPGTA-CDIOTH-MTNNGKSAAMAPVEVFGSNSEKC 210
 Db 249 IHRDKSDMLLSADKSI-KIADFGVARIYEVQTEGMTPEGTGRWMAPEMIOHRAVYNQV 307
 QY 211 DVFSGILIMETITRRKPPDEIGG--PAFRIMAVYNGRPPILKLPPISSLMRCS 268
 Db 308 DVSFGIVTWEITGLLPQNMNTAVQAFAV--VNRGRPVPVNDCLPVLSDIMTRCWD 364
 QY 269 KDPQSPSPMEELVKIM 284
 Db 365 ANPEVRPCFVEVYVKKL 380

 RESULT 7
 T10671
 protein kinase homolog F6E21.90 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cross)
 C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000

C:Accession: T10671
R:Bayan, M.; Linnard, N.; Ouail, M.; Harris, B.; Rajadream, M.A.; Barrell, B.G.; Bancroft, J.; et al.
submitted to the Protein Sequence Database, June 1999

A:Reference number: Z16533
A:Accession: T10671
A:Molecule type: DNA
A:Residues: 1-412 <BEV>
A:Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21..90
A:Experimental source: cultivar Columbia; BAC clone F6E21
C:Genetics:
A:Gene: ATSP:F6E21..90
A:Map position: 4
A:Introns: 300/2
C:Superfamily: kinase-related transforming protein; protein kinase homology F_135-352/Domains: protein kinase homology <KIN>

Query Match 14.4%; Score 434; DB 2; Length 412;
Best Local Similarity 36.8%; Pred. No. 1.1e-12;
Matches 103; Conservative 56; Mismatches 89; Indels 32; Gaps 11;

OY 25 VLNFE--IDYEIEVEEVGRGAFGVCCAKMRAKDVAIKQIE-----SESERAF 74
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db LVNYEEWTIDIRKLHMGPAFQAQFGKLYRGTYAGEDVAIKILERSDSNPKAQLAEQOF 177
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 75 IVELRLSRVNHPIIVKLYGACLNPV--CLVMEXAEGSLNVL---HGAEPLPYTAA 128
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 178 QOEVSMLFLGHPIINVIRIGACIKPMWCIVTEYAKGSVFQFLTQRNRPVLKL---- 233
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 129 HAMSWCLQCSGVAYLHSMQPKALIHRDLKPNNLLVAGTVLKICDPFTA-CDIOTH-M 186
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 234 -AVMQALDVARGMAVHE--RNFIHRDLKSDNLLISADRSI-KIADGVARIETQESGM 288
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 187 TNNGSAAAMAPEVPFEGSNYSEKCDVPSWGIIIMEVITRRKPFDEIG--PAFRIMAVH 244
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 289 TPETGYTFWMAPEMIQHPRPTOKVDVYSFGIVLWELIGLPQNMTAVQAFAV---VN 345
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 245 NGTRPPLIKNLKPRESLMTRCMSKDPQRSMEIYVKIM 284
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 346 RGVRFYTPADCPLVLGEIMTRCMDADPEVRPCFAEIVNLL 385
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 8
T18287
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (*Dictyostelium discoideum*)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C:Accession: T18287
R:Adler, K.
submitted to the EMBL Data Library, July 1996

A:Reference number: Z18856
A:Accession: T18287
A:Status: preliminary; translated from GB/EMBL/DDBU
A:Molecule type: DNA
A:Residues: 1-1338 <ADL>
A:Cross-references: EMBL:U64830; NID:g1468982; PID:g1468983; PION:AAB04999.1
C:Genetics:
A:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 14.4%; Score 433.5; DB 2; Length 1338;
Best Local Similarity 34.7%; Pred. No. 3.9e-12;
Matches 104; Conservative 54; Mismatches 113; Indels 29; Gaps 8;

OY 29 BEIDYKEIEVEEVGRGAFGVCCAKMRAKDVAIKQISE--SERKAFIVELRLQLSRVNH 86
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 1050 KEIKDEVAIYIERVGAGSFANVSLGIWNGYKVALIKLNESISNDKERIKYVSSLIKSHH 1109
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 87 ENIVKLYGACTNPVCVLMEVYAGGSLYNVH---GAEPPLPYTAHAMSWCLQCSQVYA 142
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Db 1110 PNVTWFMGARIDPPCIFTEYVLCGGSLYDVLHIQIKILNPLMMYKMHDL-----SLGME 1163
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

OY 143 YLHSMQPKALIHRDLKPNNLLVAGTVLKICDPFTACDIQTHNT-TNNGSAAAMAPPEVF 201
:::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

Matches 98 Conservative 47; Mismatches 100; Indels 23; Gaps 8;

```
OY      35 ELEVEEVRGAGFVVCCKAKRKADVAIK--QIESSEKATIVELRQ---LSRVNHPN 88  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     2 ELTLEEIIIGIGFGKKYYRAFWIGDEVAVCAARHDPDEDISOTIEVROEAKLPAALKHPN 61  
  
OY      89 IYKLVGACINP--VCLIMEVYAEGSLIYNVLHGAEPLPYTAHAAMSWCLOCSOGVAYHS 146  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     62 IIALKGVCILKEPNLCIMLFARAGGLINRLVLSGRRIIPDI---LVNNAVQLARGNNYLHD 117  
  
OY      147 MQPKALLHRDLKPENLLLV-----AGGTVLKICDFGTACD-IQTHMTNNKGSAAAMP 198  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     118 EAIVPIRHLDLKSSNILLIQKVENGNDLSNKILKITDFTGLAREMHRHTTKMSAAGTAYMAP 177  
  
OY      199 EVFEESGNYSKKCDVPSWGIILMEVITRRKPEDEIGAPARIMNAV-HNGTRPPLIKLPK 257  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     178 EVIRASMFSGSDVMSYGVLMLLELLTGVEVFRGIDG--LRVAVGANMKALPIPTCPDE 235  
  
OY      258 PIESLMTRCMSKDPSPORSPMEELVKIMT 285  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     236 PRAKLMECWNPDPHRSRPFSTINILDQLT 263
```

RESULT 12
T16747

hypothetical protein R13f6.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C/Accession: T16747
R/Miller, N.
submitted to the EMBL Data Library, April 1994
A/Description: The sequence of C. elegans cosmid R13f6.
A/Reference number: Z18570
A/Accession: T16747
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-328 <MUL>
A/Cross-references: EMBL:U00046, NID:g470358, PID:g470364, PIDN:AAC47047.1; GSPPDE:GN0002
A/Experimental source: strain Bristol N2, clone R13f6
C/Gene(s):
A:Gene: CESP.R13f6.7
A:Map position: 3
A:Introns: 20/3; 160/3; 222/2; 286/2
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h

Query Match 13.8%; Score 416; DB 2; Length 328;
Best Local Similarity 33.4%; Pred. No.5.5e-12;
Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;

```
OY      1 MSTASAASSSSSSSAGEMTIAPSQVLNFBEIDYKEILEVEEVGRCAGFVVCCKAKRKADV 60  
       ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     22 MSTPTSNESTSSSSSNNS----DQRVLPFDIORDDIOVDGHIGVGFVAVSGNWTLTPDG 76  
  
OY      61 AIKOIESESERKAFIVELRQ--LSRVNHPNIYKLGAACL--NPVCLMEVYAEGSGSLNVN 115  
       ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     77 SORTI---ALKVFLEKEAEILSKIRHNNTIIQFYGICKAGNDPFIYTEAEKSGSLYP 133  
  
OY      116 LHGAEPLEPYTTAAHAMS-----WCLQCSOGVAYHSNQKALLHRDLKPNLLVLAGGT 169  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     134 IHSESSQSAPASSGGNSPDVVVKMASQIASGIQHYDAVDITIIHRDLKSKNVIL-DKNL 192  
  
OY      170 VLKICDFGTACDIQTHMTNNK--GSAAMMAPE-VFEGSNYSKKCDVPSWGIILMEVITR 225  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     193 VKCIDDFGTSKDL-THSCAPASMGCTAAMSPSEMILQSEGGLTTATDVMSYGVVLMETLSK 251  
  
OY      226 RKFPDELGPARRIMNAV-HNGTRPPLIKLPKPIESLMTRCMSKDPSQRSMEEI 280  
       |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|  
DB     252 EVVPKYDS--EFRIPTMITQSGITLAIPSPCAPLIKQMSNCWKVTPXDRANMRQI 305
```

RESULT 13
B35670

protein-tyrosine kinase (EC 2.7.1.112) 2 - slime mold (Dictyostelium discoideum) (fragment)

C:\Species: Dictyostelium discoideum
C:\Date: 28-Sep-1990 #sequence_revision 28-Sep-1990 #text_change 04-Feb-2000
C:\Accession: B35670
R:\Tan, J.L.: Spudich, J.A.
Mol. Cell. Biol. 10, 3578-3583, 1990
A:\Title: Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoid
A:\Reference number: A35670; MUID:90287147; PMID:1972546
A:\Accession: B35670
A:\Status: preliminary
A:\Molecule type: mRNA
A:\Residues: 1-410 <TIN>
A:\Cross-references: GB:NM3784; NID:g167777; PIDN:AAA3203.1; PID:g167778
C:\Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:\Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
F:\123-371/Domain: protein kinase homology <KIN>
F:\11-122/Region: protein kinase ATP-binding motif

Query Match 13.7%; Score 414; DB 2; Length 410;
Best Local Similarity 13.5%; Pred. No. 8.4e-12;
Matches 104; Conservative 62; Mismatches 124; Indels 40; Gaps 12;

OY 28 FEE-----IDYEIEVEEVGGAFAVGVCAKPAKDVAIKOI----ESESERK 72
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 89 FEELKSIIIGEREVIITINDIOFIQKVGEAFSEWEMWGHHAIKKLTIIGDEBPK 148
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
OY 73 AFIVELROLSSVNHPNIVKLYGACLNPCVLMEYAEGSLYNVLIH--GAEPPLPYTTAAH 129
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 149 RFIREVONLKRGKHQNITVMFIGACYKPACITTEYMAGSGLYNIILHNPSSTPKVKYSFPL 208
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
OY 130 AMSWLCQSOGVAVLHSMOPKALIHRLDKPPNLLLVAGGYVLKICDPTACD--IQCHMT 187
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 209 VLRKATDMALLDLHLHSI---TVHRPLTSQNILLDDELGNI-KISDFGLSAEKSREGSMT 264
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
OY 188 NNKG---SAAWMAAEVEEG-SNVSEKCDVSFGIIMLEVITRRKPFDEIGAPFRIMAV 243
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 265 MTNGGICNPRRRPPELIRKNLGHYSEKVDVYCFSLVWEIILTGETPESDDLG-SGRASOVA 323
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
OY 244 HNGTRPPLIKNLKPIESLMTRCSKDPSPSPNEIYIKITHTLMRYFPADPELOYPCQ 303
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 324 YAGIRPRPIPEXCDELCKILLTLQCWEADPNDRPPTYIVNKLKEI-----SMNPDIGP--- 375
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
OY 304 YSDGGSNATSSTGSPMDIASTNTSKSDT 333
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:
Db 376 VSDQFYQISEPTSR---LALSNSGNSNSS 402
||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|: ||| :|:

RESULT 14
JC2363
protein kinase (EC 2.7.1.37) ZPK - human
N:\Alternate names: leucine-zipper protein kinase
C:\Species: Homo sapiens (man)
C:\Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #ext_change 02-Feb-2001
C:\Accession: JC2363
R:\Reddy, U.R.: Pleasure, D.
Biochem. Biophys. Res. Commun. 202, 613-620, 1994
A:\Title: Cloning of a novel putative protein kinase having a leucine zipper domain from
A:\Reference number: JC2363; MUID:94311945; PMID:8037767
A:\Accession: JC2363
A:\Molecule type: mRNA
A:\Residues: 1-668 <RED>
A:\Cross-references: EMBL:U07358
A:\Experimental source: brain
A:\Note: The nucleotide sequence for this amino acid sequence is inconsistent with that
he codon ACC for residue 661 as Pro, the codon GAACCACTTCACA for residues 664-668 as /
C:\Comment: This protein belongs to the family of non-receptor kinase.
C:\Genetics:
A:\Gene: GDB:ZPK
A:\Cross-references: GDB:383963; OMIM:600447
A:\Map position: 12q13-12q13
C:\Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:\Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
F:\123-371/Domain: protein kinase homology <KIN>
F:\11-122/Region: protein kinase ATP-binding motif

F:443-471/Region: leucine zipper motif
F:538-545/Region: nucleotide-binding motif A (P-loop)
F:152/Active site: Lys #status predicted

```
Query Match      13.7%; Score 412; DB 2; Length 668;
Best Local Similarity 24.6%; Pred. No. 1.7e-11;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEVGRGAFGVVCKAKWRADVAIKQIESERKAFIVELRQLSRVNHNPNI 89
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 119 EVPFEILDLQWGSQAQAVFLGRHGBEAVKVRDLKE-----TDIKHLRLKHPNI 173
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 90 VKUYGACLNLP--VCLVMEVAEGSLNVVHGAEPPLPYTAAHAMSKWCLCCSGVAYLHSM 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 174 ITFKGVCTQAPCYCIILMEFCAQQLVEVLRAGRPV---TPSLLDVMSMGAGMNYLHLH 230
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 148 QPKALHRLDKPNLLVAGGTVLKTCDFGTACDIQTHMTNNK--GSAAMWAPVPEGSN 205
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 K---IHRDLKSPN-WLITYDDVVVKISDFGTSKLSKDSKTSKWSFAGTVAWMAPEVIRNEP 286
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 206 YSEKCDVFWGIIWBEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 287 VSEKVDIWSFGVVLWELLTCEIPYKDVDSAA--IINGVGSNSLHLPVPSSCPDGFKILLR 344
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 265 RWSKDPQSQRPSMEEIVKIMTHLMRYFPGADBEPLQYPCQYSDGSGNSATSTGSMFDIAS 324
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 QCWNSKPRNRPSPRQ---ILLHL-----DIAS 368
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 325 TNT-SNKSOT-----NMEQVPATNDTIKRLSKLL----- 353
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 369 ADVLSTPQETFYFQSQAEWREVKLHFKEIKSGTCLHRLLEBELVMRRRELRLHALDIREH 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 354 -----KNQAKQOQESGRL-----SLGASHGSSVES 378
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 429 YERKLERANNLYMELNALMLQLEKRELLRREQALERCPLGKHPSPRGLLHGNTMEK 488
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 379 L-----PPTSEGRMSADMSIEARI-----AATTNGQPRRRSIQDLTVTGTGEGQVSS 428
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 489 LIKKRVNPQNLSPHSQRPDILKAESLLPKDLAALSGLVLP-----GCPKAPSPGR--S 540
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 429 RSSSPSVRMITTSQPTSEKP---TRSHPTPDSTDTNG-----SDNSIPMAYLTLDHQ 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 541 RRGKTRHRKASAKSGCDLPLGLRTVPPHPGPGSPGLGGGPSAWAECPPALRLGLHHD 600
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 480 L 480
Db 601 L 601
```

RESULT 15

F96763
hypothetical protein F25P22.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96763
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: F96763
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1030 <STO>
A:Cross-references: GB:AB005173; NID:g6692730; PIDN:AAF24836.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25P22.8

A:Map position: 1

```
Query Match      13.7%; Score 412; DB 2; Length 1030;
Best Local Similarity 32.8%; Pred. No. 2.7e-11;
Matches 104; Conservative 57; Mismatches 128; Indels 28; Gaps 10;

QY 1 MSTAASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVGRGAFGVVCKAKWRADKV 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 721 ISDKSIGNSSKSCDDVSDC-----EILWEEITVGERIGLGSYGEVYRGDWHGTEV 772
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AIKQIESE-----SERKAFIVELRQLSRVNHNPNIKLYGACLNLP--VCLVMEVAEGGSLYN 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 773 AVKFLQDDLTGEALBEFRSEVRIMKLRHPNIVLFMGAVTRPPNLSIVTEFLPRGSLYR 832
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 115 VLGHAEPPLPYTAAHAMSKWCLCCSGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKIC 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 833 LIH--RPNQLDERRLRMALDAARGMNYLHSCNP-MIVHRLKSPN-LLVDKNWVVVKVC 888
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 175 DFGTACDIQTHMT-----NNKGAAMWAPVPEGSNSYSEKCDVFWGIIWBEVITRRKPF 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 889 DFGLS--RMKHSTYLSKSTAGTAENWMAPEVLRNEPADEKCDVSYGVILWELFTLQQPW 946
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 230 DEIGGPAFRIMWAVHNGTRPPLIKNLPKP-IESIMTRCWSKDPQSQRSMEEIVKIMTHLM 288
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 947 GKMN--PMQVVGAVGFOHRLDIPDFVDPALADLISKWOTDSKLRPSFAEIMASLKRLO 1004
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 289 RYFPGADEPLQYPCQYS 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1005 KPVTSNIPRPVSSSS 1021
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

Search completed: December 9, 2002, 22:59:38
Job time : 38.1468 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model
Run on: December 9, 2002, 22:45:29 ; Search time 62.5512 Seconds
(without alignments)
1907.257 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAGSSSSSSAGEMIE.....QCKKQLEVRISQQKROGTS 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : SPTREMBL_21.*
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriap.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2963.5	98.3	606	11 Q923A8	Q923A8 mus musculus
2	2819	93.5	539	4 Q9NTR1	Q9NTR1 homo sapien
3	2795.5	92.8	566	4 Q9NTR2	Q9NTR2 homo sapien
4	2638.5	87.5	616	13 Q73613	Q73613 xenopus lae
5	2512	83.3	491	4 Q9NZ70	Q9NZ70 homo sapien
6	2301.5	76.4	478	4 Q9NTR4	Q9NTR4 homo sapien
7	1310	43.5	260	4 Q9UG54	Q9UG54 homo sapien
8	966	32.1	678	5 Q9VJ06	Q9VJ06 drosophila
9	505	16.8	252	5 Q9VCV0	Q9VCV0 drosophila
10	481	16.0	455	4 Q9HCC4	Q9HCC4 homo sapien
11	478	15.9	454	11 Q9ESL3	Q9ESL3 mus musculus
12	462.5	15.3	1066	4 Q9H2N5	Q9H2N5 homo sapien
13	461.5	15.2	850	11 Q9JJI5	Q9JJI5 mus musculus
14	459.5	15.2	847	4 Q16584	Q16584 homo sapien
15	457	15.2	1001	11 Q8VDG6	Q8VDG6 mus musculus
16	456.5	15.1	800	4 Q9HDD2	Q9HDD2 homo sapien

17	456.5	15.1	800	4 Q9HCC5	Q9HCC5 homo sapien
18	456.5	15.1	800	4 Q9NYL2	Q9NYL2 homo sapien
19	456.5	15.1	800	4 Q9NYE9	Q9NYE9 homo sapien
20	456.5	15.1	1036	4 Q8WNN1	Q8WNN1 homo sapien
21	454.5	15.1	802	11 Q9ESL4	Q9ESL4 mus musculus
22	452.5	15.0	536	5 Q9XTC6	Q9XTC6 caenorhabdi
23	447.5	14.8	371	13 Q90ZY8	Q90ZY8 brachydanio
24	446	14.8	564	4 Q9HIY7	Q9HIY7 homo sapien
25	442	14.7	520	4 Q8WNN2	Q8WNN2 homo sapien
26	436.5	14.5	1020	5 Q9W3I3	Q9W3I3 drosophila
27	436.5	14.5	1148	5 Q9SVF6	Q9SVF6 drosophila
28	436.5	14.5	1161	5 Q95UN8	Q95UN8 drosophila
29	434.5	14.4	411	10 Q9ZQ31	Q9ZQ31 arabisdopsis
30	434	14.4	412	10 Q9M085	Q9M085 arabisdopsis
31	433.5	14.4	422	5 Q23846	Q23846 dictyosteli
32	433.5	14.4	1338	5 Q23927	Q23927 dictyosteli
33	427.5	14.2	391	10 Q22100	Q22100 arabisdopsis
34	424	14.1	977	5 Q9VW24	Q9VW24 drosophila
35	421.5	14.0	357	10 Q9SIM8	Q9SIM8 arabisdopsis
36	419.5	13.9	527	5 Q9BI25	Q9BI25 dictyosteli
37	418.5	13.9	416	10 Q94C42	Q94C42 triticum ae
38	418	13.9	835	4 Q9Y2V6	Q9Y2V6 homo sapien
39	416	13.8	740	5 Q21982	Q21982 caenorhabdi
40	412	13.7	483	10 Q8RY96	Q8RY96 arabisdopsis
41	412	13.7	570	10 Q8RWL6	Q8RWL6 arabisdopsis
42	412	13.7	1030	10 Q9C9U5	Q9C9U5 arabisdopsis
43	411.5	13.7	546	10 Q22558	Q22558 arabisdopsis
44	411.5	13.7	553	10 Q81808	Q81808 arabisdopsis
45	411	13.6	859	4 Q8WY25	Q8WY25 homo sapien

ALIGNMENTS

RESULT 1

ID Q923A8 PRELIMINARY; PRT; 606 AA.

AC Q923A8; 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Unknown (protein for MGC:5989).

GN MAP3K7.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC006665; AA06665.1; -.

DR MGD; MGI:1348877; Map3k7.

DR InterPro; IPR000719; Euk_pkinase.

DR InterPro; IPR002290; Ser_thr_pkinase.

DR Pfam; PF00069; pkinase; 1.

DR ProDom; PD000001; Euk_pkinase; 1.

DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.

KW ATP-binding; Transferase.

SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Query Match 98.3%; Score 2963.5; DB 11; Length 606;
Best Local Similarity 94.7%; Pred. No. 2.4e-193;
Matches 574; Conservative 1; Mismatches 4; Indels 27; Gaps 1;

QY 1 MSTASAGSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKV 60
|||||

Db 1 MSTASAGSSSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKRAKV 60
|||||

QY 61 ATKQESSERKAFIVELRQSRVNHPIVKYGLNPVCLVMEYAGGSLYNVLHAE 120
|||||

DB 61 AIKQIESESRKAFIYELRQLSRVNHNIYKLYGACLNIPVCLVMEYAEGLYNNVHGA 120
 QY 121 PLEPYTAAHAMSWCLOCSQGVAYLHSMQKALLIHRDLKPNLLLVAGTVLKICDFGTAC 180
 DB 121 PLEPYTAAHAMSWCLOCSQGVAYLHSMQKALLIHRDLKPNLLLVAGTVLKICDFGTAC 180
 QY 181 DICTHMTNNKGSAAAMAPAEVFEBSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 DB 181 DICTHMTNNKGSAAAMAPAEVFEBSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
 QY 241 MAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEIVKIMTLMRYFPGADEPLQY 300
 DB 241 MAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEIVKIMTLMRYFPGADEPLQY 300
 QY 301 PCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
 DB 301 PCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNQAKQ 360
 QY 361 SESGRSLGASHGSSVESLPPTSEGRMSADMEIEARIAT----- 402
 DB 361 SESGRSLGASHGSSVESLPPTSEGRMSADMEIEARIATVAAYSKPRGRKTAISFGN 420
 QY 403 -----TGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPTRSH 453
 DB 421 ILDVPEIVISGNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPARSH 480
 QY 454 WTPDDSTDTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQYMKVQTEI 513
 DB 481 WTPDDSTDTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQYMKVQTEI 540
 QY 514 ALLQKQELVAELDDEKQONTSLVQEHKKLLDENKSLSTYYQCKKQLEVIRSOQ 573
 DB 541 ALLQKQELVAELDDEKQONTSLVQEHKKLLDENKSLSTYYQCKKQLEVIRSOQ 600
 QY 574 KROGTS 579
 DB 601 KROGTS 606

RESULT 2
 Q9NTR1 PRELIMINARY; PRT; 539 AA.
 ID Q9NTR1;
 AC Q9NTR1;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE DJ154G14.1.1 (Mitogen-activated protein kinase kinase 7
 DE (TGF-beta activated kinase 1a (TAK1)) (Fragment).
 GN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87607.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 FT NON_TER 1

SQ SEQUENCE 539 AA; 6006 MW; E6183F553CC7F324 CRC64;
 Query Match 93.5%; Score 2819; DB 4; Length 539;
 Best Local Similarity 99.8%; Pred. No. 1,4e-183;
 Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 41 VVGKAGFVGVCKAKWRAKDVAIKQIESESRKAFIYELRQLSRVNHNIYKLYGACLNIPV 100
 DB 1 VVGKAGFVGVCKAKWRAKDVAIKQIESESRKAFIYELRQLSRVNHNIYKLYGACLNIPV 60
 QY 101 CLVMEYAEGLSLVNVHGAEPYPTAAHAMSWCLOCSQGVAYLHSMQKALLIHRDLKPP 160
 DB 61 CLVMEYAEGLSLVNVHGAEPYPTAAHAMSWCLOCSQGVAYLHSMQKALLIHRDLKPP 120
 QY 161 NLLVAGTVLKICDFGTACDIQTHMTNNKGSAAAMAPAEVFEBSNYSKCDVFSWGIILW 220
 DB 121 NLLVAGTVLKICDFGTACDIQTHMTNNKGSAAAMAPAEVFEBSNYSKCDVFSWGIILW 180
 QY 221 EVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEI 280
 DB 181 EVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNLPKPIESLMTRCMSKDPQSPSMEEI 240
 QY 281 VKIMTHLMRYFPGADEPLQYPCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPA 340
 DB 241 VKIMTHLMRYFPGADEPLQYPCQYSDGQNSATSTGSEFMDIASTNTSNKSDTNMEQVPA 300
 QY 341 TNDTIKRLSKLLKNQAKQOESGRSLGASHGSSVESLPPTSEGRMSADMEIEARIA 400
 DB 301 TNDTIKRLSKLLKNQAKQOESGRSLGASHGSSVESLPPTSEGRMSADMEIEARIA 360
 QY 401 ATTNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPTRSHPTDDST 460
 DB 361 ATTNGQPRRRSIODLTVTGTEPGQVSSRSSSPSVRMITTSGPTSEKPTRSHPTDDST 420
 QY 461 DTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQYMKVQTEIALLQK 520
 DB 421 DTNGSDNSIPMAVYLLDHOLOPLAPCNSKESMAVFEQCKMAQYMKVQTEIALLQK 480
 QY 521 QELVAELDDEKQONTSLVQEHKKLLDENKSLSTYYQCKKQLEVIRSOQKROGTS 579
 DB 481 QELVAELDDEKQONTSLVQEHKKLLDENKSLSTYYQCKKQLEVIRSOQKROGTS 539

RESULT 3
 Q9NTR2 PRELIMINARY; PRT; 566 AA.
 ID Q9NTR2;
 AC Q9NTR2;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE DJ154G14.1.2 (Mitogen-activated protein kinase kinase 7
 DE (TGF-beta activated kinase 1b (TAK1)) (Fragment).
 GN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -i- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87606.1; -.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC_1
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE; PS00101; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW KINASE; Serine/threonine-protein kinase.
FT NON TER 1
SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;

Query Match 92.8%; Score 2795.5; DB 4; Length 566;
Best Local Similarity 95.1%; Pred. No. 5.8e-182;
Matches 538; Conservative 0; Mismatches 1; Indels 27; Gaps 1;

QY 41 VVGRGAFGVCKAKWRAKDVAKQIESERKAFIVELROLSRVNHNIKLYGACLNVP 100
Db 1 VVGRGAFGVCKAKWRAKDVAKQIESERKAFIVELROLSRVNHNIKLYGACLNVP 60

QY 101 CLVMEYAEGLSYLVNHLHGAELPYTTAAHMSWCLQSQGVAYLHSMQPKALIHRLDKPP 160
Db 61 CLVMEYAEGLSYLVNHLHGAELPYTTAAHMSWCLQSQGVAYLHSMQPKALIHRLDKPP 120

QY 161 NLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAWMAPEVFEKCDVFSWGIILW 220
Db 121 NLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAWMAPEVFEKCDVFSWGIILW 180

QY 221 EVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNLPKPIESLMTKRCWSKDPSPRPSMEI 280
Db 181 EVITRRKPFDEIGGPAFRIMWVHNGTRPPLIKNLPKPIESLMTKRCWSKDPSPRPSMEI 240

QY 281 VKIMTHLMRYPPGGADEPLQPCQYSDGQSNATSTGSMFMDIASTNTSNKSDTNMQVPA 340
Db 241 VKIMTHLMRYPPGGADEPLQPCQYSDGQSNATSTGSMFMDIASTNTSNKSDTNMQVPA 300

QY 341 TNDTIKLESKLLKNQAKQOSESRLSLGASHGSSVESLPTSEGKRMASDMSEIEARIA 400
Db 301 TNDTIKLESKLLKNQAKQOSESRLSLGASHGSSVESLPTSEGKRMASDMSEIEARIA 360

QY 401 ATT-----GNGQPRRSIQDLTVGTGTEPGQVSSRSSSP 433
Db 361 ATTAYSFKPKGHRKTASFGNILDVPEIVISGNGQPRRSIQDLTVGTGTEPGQVSSRSSSP 420

QY 434 SVRMITTSPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHLOLQPLAPCPNSKESM 493
Db 421 SVRMITTSPTSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHLOLQPLAPCPNSKESM 480

QY 494 AVPEQHCMAQYMKVQTEIALLQKQELVAELDQDEKQOQNTSLRVQEHKKLLDENKS 553
Db 481 AVPEQHCMAQYMKVQTEIALLQKQELVAELDQDEKQOQNTSLRVQEHKKLLDENKS 540

QY 554 LSTYVQCKQKQLEVIRSQOQKRGTS 579
Db 541 LSTYVQCKQKQLEVIRSQOQKRGTS 566

RESULT 4
O73613 PRELIMINARY; PRT; 616 AA.
AC O73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAK1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
development.";
RL EMBO J. 17:1019-1028(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.

DR EMBL; U92030; AAC14008.1; -.
DR HSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; STY_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_kinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00108; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Query Match 87.5%; Score 2638.5; DB 13; Length 616;
Best Local Similarity 84.3%; Pred. No. 3.1e-171;
Matches 519; Conservative 20; Mismatches 28; Indels 49; Gaps 7;

QY 12 SSSAGEMIEAFSQVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDVAKQIESER 71
Db 2 SATSAEMIETP-PVLNFEIDYKEIEVEEVVGRGTFGVVCKAKWRAKDVAKQIESER 60

QY 72 KAFIVELQRLSRVNHNPVIVKLYGACLNVPCLVMEYAEGLSYLVNHLHGAELPYTTAAHAM 131
Db 61 KAFIVELQRLSRVNHNPVIVKLYGACLNVPCLVMEYAEGLSYLVNHLHGAELPYTTAAHAM 120

QY 132 SWCLQSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKG 191
Db 121 SWCLQSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKG 180

QY 192 SAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPL 251
Db 181 SAAMWAPVFEKSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWVHNGTRPPL 240

QY 252 IKNLKPKPIESLMTKRCWSKDPSPRPSMEIIVKIMTHLMRYPPGGADEPLQPCQYSDGQSN 311
Db 241 IKNLKPKPIESLMTKRCWSKDPSPRPSMEIIVKIMTHLMRYPPGGADEPLQPCQYSDGQSN 300

QY 312 SATSTGSMFMDIASTNTSNKSDTNME---OVPAITNDTIKLESK---LKNQAKQOSES 364
Db 301 SATSTGSMFMDIASTNTSNKSDTNME---OVPAITNDTIKLESK---LKNQAKQOSES 360

QY 365 RLKLGASHGSSVESI-----PP-----TSEGKRMASDMSEIEARIAATT----- 403
Db 361 RLKLGASHGSSVESI-----PP-----TSEGKRMASDMSEIEARIAATT----- 420

QY 404 -----GNGQPRRSIQDLTVGTGTEPGQV---SSRSSSPSVRMITTSGP 443
Db 421 TASFGNILDVPEIITAGNGQOQRRSVQDLTVGTGTEPGQVSSRSSSPSVRMITTSGP 480

QY 444 TSEKPTRSHPTDDSTDTNGSDNSIPMAYLTLDHLOLQPLAPCPNSKESMAVFEQHCMA 503
Db 481 TDPKPRGLPWAPDESDTNGSDNSIPMAYLTLDHLOLQPLAPCPNSKESMAVFEQHCMA 540

QY 504 QEYMKVQTEIALLQKQELVAELDQDEKQOQNTSLRVQEHKKLLDENKSLSTYVQCKK 563
Db 541 QEYMKVQTEIALLQKQELVAELDQDEKQOQNTSLRVQEHKKLLDENKSLSTYVQCKK 600

QY 564 QLEVIRSQOQKRGTS 579
Db 601 QLEVIRSQOQKRGTS 616

RESULT 5
Q9NZ70 PRELIMINARY; PRT; 491 AA.
ID Q9NZ70
AC Q9NZ70; Q9NTR3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE TGF beta-activated kinase splice variant d (DJ154G1.1.4)
DE (mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated
DE kinase 1d (TAK1))).

GN TAK1 OR MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=2056765; PubMed=1118615;
 RA Demsey C.E., Sakurai H., Sugita T., Guesdon P.;
 RT "Alternative splicing and gene structure of the transforming growth
 factor beta-activated kinase 1.";
 RL Biochim. Biophys. Acta 1517:46-52(2000).
 RN [2]
 RP SEQUENCE OF 41-491 FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AF218074; AAF27652.1; -;
 DR EMBL; AL121964; CAB87605.1; -;
 DR HSSP; P08631; 1AD5.
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 491 AA; 53739 MW; B7D8832E286A59C5 CRC64;

Query Match 83.3%; Score 2512; DB 4; Length 491;
 Best Local Similarity 99.0%; Pred. No. 9e-163;
 Matches 481; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 MSTRASASSSSSSAGMIFAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKAKMRKADV 60
 DB 1 MSTRASASSSSSSAGMIFAPSOVLNFEIDYKEIEVEEVVGRGAFVCAKAKMRKADV 60
 QY 61 AIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPVCLMEYAEAGSLYNVLHGAE 120
 DB 61 AIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPVCLMEYAEAGSLYNVLHGAE 120
 QY 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGVLYKICDEGTAG 180
 DB 121 PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGVLYKICDEGTAG 180
 QY 181 DIQTHMNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILMEVITRRKKPFDELGGPAFRIM 240
 DB 181 DIQTHMNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILMEVITRRKKPFDELGGPAFRIM 240
 QY 241 WYVHNGTRPPLIKNLPRPISLMTKRCWSKQPSQPSMEELVKIMTHLMRYFPADDEPLY 300
 DB 241 WYVHNGTRPPLIKNLPRPISLMTKRCWSKQPSQPSMEELVKIMTHLMRYFPADDEPLY 300
 QY 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIRLESKLLKNQAKOO 360
 DB 301 PCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPAINTDTIRLESKLLKNQAKOO 360
 QY 361 SESGRLLGASHGSSVESLPTSECKMSADMEIEARIAATGNGQPRRSIIDLTVTG 420
 DB 361 SESGRLLGASHGSSVESLPTSECKMSADMEIEARIAATGNGQPRRSIIDLTVTG 420
 QY 421 TEPCQVSSRSSSPSVNMTTSGPTSEKPTRSHPTPDSTDTNGSDNSIPMAYLLTDHOL 480
 DB 421 TEPCQVSSRSSSPSVNMTTSGPTSEKPTRSHPTPDSTDTNGSDNSIPMAYLLTDHOL 480
 QY 481 QPLAPC 486

DB 481 QARTSC 486

RESULT 6
 Q9NTR4 PRELIMINARY; PRT; 478 AA.
 ID Q9NTR4;
 AC Q9NTR4;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE D154G14.1.3 (Mitogen-activated protein kinase kinase 7
 DE (TGF-beta activated kinase 1c (TAK1))) (Fragment).
 GN MAP3K7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tracey A.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87604.1; -;
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase; Serine/threonine-protein kinase.
 FT NOW TER
 SQ SEQUENCE 478 AA; 52482 MW; 177CC8CBA8D8DF8 CRC64;

Query Match 76.4%; Score 2301.5; DB 4; Length 478;
 Best Local Similarity 93.2%; Pred. No. 1.8e-148;
 Matches 441; Conservative 0; Mismatches 5; Indels 27; Gaps 1;

QY 41 VVGGAEGVVCWKAKRKDAIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPV 100
 DB 1 VVGGAEGVVCWKAKRKDAIKQIESESEKKA FIVELRQLSRVNHPIVLYGACINPV 100
 QY 101 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPP 160
 DB 101 CLVMEYAEAGSLYNVLHGAEPLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPP 160
 QY 161 NILVAGGTVLKICDFGTACDIQTHMNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILW 220
 DB 161 NILVAGGTVLKICDFGTACDIQTHMNNKGSAAAMAPAEVFECSNYSKCDVFSWGIILW 220
 QY 221 EYITRRKPFDEIGGAPFIMAVHNGTRPPLIKNLPRPISLMTKRCWSKQPSQPSMEEL 280
 DB 221 EYITRRKPFDEIGGAPFIMAVHNGTRPPLIKNLPRPISLMTKRCWSKQPSQPSMEEL 280
 QY 281 VKIMTHLMRYFPADDEPLYPCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPA 340
 DB 281 VKIMTHLMRYFPADDEPLYPCQYSDGQSNATSTGSPMDIASTNTSNKSDTNMEQVPA 340
 QY 341 TNDTIRLESKLLKNQAKQESGRSLGASHGSSVESLPTSECKMSADMEIEARIA 400
 DB 341 TNDTIRLESKLLKNQAKQESGRSLGASHGSSVESLPTSECKMSADMEIEARIA 400
 QY 401 ATT-----GNGQPRRSIIDLTVTGTEPCQVSSRSSSP 433
 DB 361 ATTAVSKRGRKRTASGNILVPEIYISGNGQPRRSIIDLTVTGTEPCQVSSRSSSP 420
 QY 434 SVRMITTSGPTSEKPTRSHPTPDSTDTNGSDNSIPMAYLLTDHOLQPLAPC 486

Db 421 SVRMITTSGTSEKPTRSHPTWPDSTDTNGSDNSIPMAYLTLDHQLQARTSC 473

RESULT 7

ID Q9UG54 PRELIMINARY; PRT; 260 AA.
AC Q9UG54;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE Hypothetical 28.9 kDa protein.
GN DKF2P566F0420.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=UTERUS;
RA Wambutt R., Heubner D., Mewes H.W., Gassenhuber J., Wiemann S.;
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL050393; CAB43687.2; -;
KW Hypothetical protein.
SQ SEQUENCE 260 AA; 28870 MW; D24F292932FB7F72 CRC64;

Query Match 43.5%; Score 1310; DB 4; Length 260;
Best Local Similarity 99.6%; Pred. No. 2e-81;
Matches 259; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 320 MDIASTNTSNKSDTNMEQVPAATNDTIKRLESKLLKNQAKQSSGRLSLGASHGSSVESL 379
|
|
|
Db 1 MDIASTNTSNKSDTNMEQVPAATNDTIKRLESKLLKNQAKQSSGRLSLGASHGSSVESL 60
|
|
|
QY 380 PPTSEKGMKMSADNSETIARAAATNGCQPRRSIQDLTVTGTGEPGVSSRSSPSVRMIT 439
|
|
|
Db 61 PPTSEKGMKMSADNSETIARAAATNGCQPRRSIQDLTVTGTGEPGVSSRSSPSVRMIT 120
|
|
|
QY 440 TSQGTSEKPTRSHPTWPDSTDTNGSDNSIPMAYLTLDHQLQPLAFCPNKSKESMAVFEGH 499
|
|
|
Db 121 TSQGTSEKPTRSHPTWPDSTDTNGSDNSIPMAYLTLDHQLQPLAFCPNKSKESMAVFEGH 180
|
|
|
QY 500 CKMAQYMKVQTEIALLQKQELVAELDQDKQNTSLRQVHKLLDENKSLSTYYQ 559
|
|
|
Db 181 CKMAQYMKVQTEIALLQKQELVAELDQDKQNTSLRQVHKLLDENKSLSTYYQ 240
|
|
|
QY 560 QCKKQLEVIQSQQKRGTS 579
|
|
|
Db 241 QCKKQLEVIQSQQKRGTS 260
|
|
|

RESULT 8
QYV306 PRELIMINARY; PRT; 678 AA.
ID QYV306
AC QYV306;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative TAK1 protein (CG1388 protein) (LD42274P).
GN TAK1 OR CG1388 OR CG18492.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pharyngota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Cocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.G., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.P., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.-N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
RA O'Connor M.B., Shibuya H., Ueno N.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003571; AAF50895.1; -;
DR EMBL; AF199466; AAF06815.1; -;
DR EMBL; AY051953; AAK93377.1; -;
DR HSSP; P08631; IAD5
DR Flybase; FBgn026323; Tak1.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001990; Granin.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00422; GRANINS_1; UNKNOWN 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8CDE45 CRC64;

Query Match 32.1%; Score 966; DB 5; Length 678;
Best Local Similarity 36.0%; Pred. No. 1.8e-57;
Matches 226; Conservative 97; Mismatches 207; Indels 98; Gaps 14;

```

QY 18 MIFAPSOVLNFEIDYKEIEVEEVNGARGVYCKAKWRAKQVAIKQISESEKRAPIVE 77
DB 1 MATASLDALAAAYVDFEILREKVGSGYGVCKAWRKLAIVAEFFASAKQKIEKE 60
QY 78 LROLSRVNHNPIVLYG--ACLNPCVLMVEAEGSLVYVHGAEPYPYTAHANSMWCL 135
DB 61 VKOLSRKHNNIILAHIOISSYQATYLIMEFAGGSIAHNLHG-KYKPAYSLAHAMSMAR 119
QY 136 QCSQGAVALYHSOPKALIHBDLKPILLVAGSTVLKICDFGACDICTQHTMNNKSAAM 195
DB 120 QCAGEGLAYLHAMTPKPLHRDVKPMLLLTNKGRNLKICDFGVAADKSTMTMTNNGSAAM 179
QY 196 MAEVEPVGSGVSKCDVFSGIIIMVEITRRKPFDELGGAFPIMAVNGRRPLIKXL 255
DB 180 MAEVEPVGSGVSKCDVFSMAIIVLVEVLSKOPFGIDN-AVTTQWKIYKGERPPLTTCC 238
QY 256 PKDIESIMTRCWSKDSQSPSMEEIVKIMTHLMRYFPGADEPQY----- 300
DB 239 PKRIEDLMTCKMTVPEDRPSMQYIGVNHIEYKDYTGADKALETFVVAQIYTKESDGT 298
QY 301 -----PCQYSDGEGSN-----SATSTGSEMDIASVTNTSKSDTN 334
DB 299 VAAQPSLSSQEGELSPSSSTQLTPTTAANVANVAIAISKTTTSSMTENTSSSTSDITPTN 358
QY 335 MEQVP-----ATN--DTIKLESKLKN-----QAKQSSGSLSGASHGSSV 376
DB 359 SGOLDNNPLFYVYVYVNRMDALPEEESNESRSDSNLTSSAEATQRTETTRNGMLWACKPM 418
QY 377 ESLEPTSEGRMSADMSEIARIAATGNGQPRRSIODLTVTGTEP-----GQ 425
DB 419 EQL--TLVDVANGPDLSPSSSSSSSTNAKSDGER---LTVTDTPVMMTTDLSNNGG 472
QY 426 VSSRSSPSVTRMTTGTGPTSEKPTRSHPW-----TPDSDTNGSDNSIPVAYTL 476
DB 473 IHAHNSGLSHANGQARDELOEHEOEIYNSLDVDVDPDEDNDGTEGOSIAE---IL 529
QY 477 DHLOLAPLPNPKESMAVPECHCKMAOEYMKVQTEIALLQKQELVAVELDODEKQDN 536
DB 530 DEPLQEPPTPNDASQLYIRHHRMAKEVLSVDTLVYAQDFKDLIVQMDRTEREQK- 588
QY 537 TSLVQEHKKLDEKNSLSTYYQCKKQ 564
DB 589 -----QELLRKMKDKGLOSLVNNLQQ 611

```

```

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dubin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaeser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Houtlin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishina N.V., Mobary B., Murphy L., Murry D.M., Nelson D.L.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murry D.M., Nelson D.L.,
RA Najazolo M., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palenzola M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner C., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Glabs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AEO03740; AAF56055.1;
DR FLYBASE: FBgn0039015; CG4803.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PSS00107; PROTEIN KINASE ARP; 1.
DR PROSITE: PSS0011; PROTEIN KINASE_DOM; 1.
DR PROSITE: PSS0108; PROTEIN_KINASE_ST; 1.
DR KWP: ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 252 AA; 28853 MW; 8E006F2EAB68D17 CRC64;

```

Query Match 16.8%; Score 505; DB 5; Length 252;
Best Local Similarity 40.6%; Pred. No. 1e-26;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

```

QY 29 EEIDYKEIEVEEVNGARGVYCKAKWRAKQVAIKQISESEKRAPIVELROLSRVNHPN 88
DB 6 EGVPEYERIQTKELGTGFGYSVYRAVWRNREILAKRIRECECEKXIRELYOLTASHVN 65
QY 89 IVKIYGACLPVC--LWVEYAEGLSLVYVHGAEPYPYTAHANSMWCLQCSQGAVALYHS 146
DB 66 IVELYGTSRHEGCLALIMEFVDDGSLSSFLH-AKSPSYSHAHAFWMAHQIAGIAYLHG 124
QY 147 MQPKALIHBDLKPENLLVAGSTVLKICDFGACDICTQHTMNNKGSAAAMP----- 199
DB 125 MQPKAVIHRDILKPLNLLTCEKGLKIKICDFGTVDLSQSTSCAAGCRKAPABRVLEDF 184
QY 200 -----VFEAGSYSEKCDVFSGIIIMVEITRRKPFDELGGAFPIMAVNGRRPLIKXL 244
DB 185 KSNRIINOPTGQKVLQGNKPKPEKCDVYSMAITPFEILSRKEPFQY-NTLPELYMAIN 243
QY 245 NGTR 248
DB 244 EGKR 247

```

```

RESULT 9
O9VCVO PRELIMINARY; PRT; 252 AA.
AC O9VCVO;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydroidea; Neuroptera; Embiopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceinliker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriti J.F., Adyanti A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

```



```

Db 173 MADEVIQSLSEVSECTDYSGVILWEMLTREVPFKGLEG--LQVWLVEKMERLTI PSS 230
QY 255 LKPIESLMTKRCWSKDSQSPSMEIYKINTHLMRYFPGADEPLQYPCQYSDGQSNKAT 314
Db 231 CPSPFALLHQCEWADAKKPSFKQIISLI-----ESMSNDT- 267
QY 315 STSPFMDIASTNTSNKSDTMEQVPATNDTIKRLSKSLTKNQAKQOSESGRLSL---GAS 371
Db 268 ---NLDPQCSPLANKAEMNCE-IEATLERLKLERDLSEKQELKERERRLKMEQKLT 323
QY 372 HGSSVESLPTST-----EGKMSADNSEIEARI-AAITGNGQPRRRSIQDITVTGT 421
Db 324 EQSNTPLLLPLLSARMSSESYFESKTEESNSAEMSCQITAAISNGHGGMNPGLOAMMLMGF 383
QY 422 EPGQVSSRSSSPSV 435
Db 384 --GDVFSMNKGAIV 395

```

RESULT 12

```

O9H2N5
ID O9H2N5 PRELIMINARY; PRT; 1066 AA.
AC O9H2N5;

```

```

DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA SEQUENCE FROM N.A.
RA McNea J.J., Dower S.K., Guesdon F.;
RT "cDNA sequence and gene organisation of mixed lineage kinase 1.";
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AF251442; AAG44591.1; -.
DR HSSP; P29355; ISEM.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR pfam; PF00069; pkinase; 1.
DR PRINTS; PR00018; SH3; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
DR KINASE; SH3 domain.
FT NON_TER 1
SQ SEQUENCE 1066 AA; 118463 MW; EDD08EBE7482723 CRC64;

```

```

Query Match 15.3%; Score 462.5; DB 4; Length 1066;
Best Local Similarity 27.0%; Pred No. 5,4e-23;
Matches 141; Conservative 81; Mismatches 192; Indels 109; Gaps 15;

```

```

QY 3 TASAASSSSSSSAGE--MIEAPSOVLNFEIDYKEIEVEEVVGAFGVVCAKMRADV 60
Db 60 TRSAFSSSCQGGEDPCYPIQL---EIDFAELTLEIIGIGFGVYRAFMIGDEV 116
QY 61 AIK--QIESESRKAFIVLRQ---LSRVNHPNIVKLYGACLANP--VCLVMEVYEGSL 112
Db 117 AVKARHDPDEDISQTIENVROEAKLFMLKHPNIIALRGVCLKEPNCLVMEFARGGPL 176

```

```

QY 113 YNVLHGAEPLPYTAAHAMSWCLQCSQGVAVLHSMQKALIHRLDKPPNLLV----- 165
Db 177 NRVLGKRIIPPD-----LVNMAVOIAGMNYIHDEALVPIIHRDLKSNLILIQKVENCD 232
QY 166 AGTVLTKICDFGTACD-IQTHMTNNKSAAMAEBEVEGGSNYSEKQDVFSGIILEVIT 224
Db 233 LSNKILKTIPTGLAREWHRRITTKMSAGTYAMMADEVIRASMFSGSDVMSYGVILWELLT 292
QY 225 RRRPFDELGGAFIIMAVHNGTRPPLIKNLKPIESLMTKRCWSKDSQSPSMEIYKIM 284
Db 293 GEVPPRGIDGLAVAVGYAM-NKUALPIPTCPEPFAXLAMEDCNPPHSPRSPFTNILDQL 351
QY 285 TLHMRYPGADEPLQYPCQYSDGQSNATSTGTF-----MDIASNTSNKSDTMEQV 338
Db 352 T-----TIESGFPEPKDSFHCLODNMKHEIQEMFDQL 385
QY 339 PATNDTIKRLSKSLTKNQAKQOSESGRL-----S 367
Db 386 RAKEKEI RTWEEELTRAALQCKQOEILRRREQELAREIDILEREINIIHQLOCEKPR 445
QY 368 LGASHGSSVESLPTSTSGKMSADMSIEARIATGNGQPRRRSIQDITVTGTBQVNS 427
Db 446 VKRRKGKFRKSRRLKDKGNRISLP-SDPHKFTVOASPTWDKRKSLL-----IN 492
QY 428 SRSSSPSVRMITTSQPTSEKPTRSHPTPDSTDTNGSDNSIP 470
Db 493 SRSSSP-----ASPTIIRLRAIQLTPGESSKTKWGRSSVVP 528

```

RESULT 13

```

O9JUL5
ID O9JUL5 PRELIMINARY; PRT; 850 AA.
AC O9JUL5;

```

```

DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Mixed lineage kinase 3.
GN MLK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RA SEQUENCE FROM N.A.
RA STRAIN=129/OLA.
RX MEDLINE=20354997; PubMed=10894943;
RA Saridaki A., Ferraz C., Demaille J., Scherrer G., Roux A.-F.;
RT "Genomic sequencing reveals the structure of the Kmk6 and Map3k1
RT genes and their close vicinity to the Sipa1 gene on mouse chromosome
RT 19.";
RL Cytogenet. Cell Genet. 89:85-86(2000).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AF15142; AAF73281.1; -.
DR HSSP; P29355; ISEM.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR pfam; PF00069; pkinase; 1.
DR pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50002; SH3; 1.
DR ATP-binding; Kinase; SH3 domain; Serine/threonine-protein kinase;
KW Transferase.

```


RESULT 15
Q8VDG6 PRELIMINARY; PRT; 1001 AA.
ID Q8VDG6;
AC Q8VDG6;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase 9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021891; AAH21891.1; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PRO0452; SH3DOMAIN.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; UNKNOWN_1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
DR PROSITE; PS00002; SH3; 1.
KW kinase.
SQ SEQUENCE 1001 AA; 109983 MW; E10042C68B9953C CRC64;

Query Match 15.2%; Score 457; DB 11; Length 1001;
Best Local Similarity 26.7%; Pred. No. 1.2e-22;
Matches 159; Conservative 89; Mismatches 219; Indels 128; Gaps 23;

QY 31 IDYKEIEVEEVGGAFGVVCYAKWRAKDYAIKQISESEKRFIV-----ELRQLSRV 84
DB 105 VDERERLEKELIGAGGFGVYRAITWQGEYAVAAARRDEQDAAAASVREARLFAML 164
QY 85 NHPNIVKLYGACLN--VCLVMEYAEGLYNYLHGAEPPLPYTAAH-----AMSW 133
DB 165 RHENITQLRGVCLRPHLCVLEFARGALNRALAAASDPRAPGRARATPPQVLVNW 224
QY 134 CLOCSQGVATLHSMOKALIHRLKPPNLLVAG-----GTVLKICDFTACD-IQTH 185
DB 225 AVQIARGMLYLHEAVVPIIHRDLKSSNILLLEKIEHDDICNKTLPFGIAREWHRTT 284
QY 186 MTNKKSSAAMMAPEVEFGSNVSEKCVFSGIILMEVITRRKPFDEIGGPAFRIMAVAN 245
DB 285 RMSAAGTYAMMAPEVIRRSIFSGSDIWSYGVLLMELITGEVPRGIDGLAVAYGAVVK 344
QY 246 GTRPPLIKLKPPIESLMTRCWSKDSQPSMEIVKIMTHL-----MRYPGADE 296
DB 345 LTL-PIPTCTPEPFATLKECWEQDPIRPSFALLIQQLTAIEAVLTMWQESFHSQE 403
QY 297 PLOYPQYSDSEGQSNATSTGSMFMDIASTNTSKSD---TNMEQVPATN--DTIKRLS 350
DB 404 DWLLEIOMSESLRTKEKELRSREELSRALQKSOELLRRREQLAREITDVLER-EL 462
QY 351 KLLKNOAKOOS-----ESGRSLGSHSHSVES-----LPPTSEGRMSAD 391
DB 463 NVLIFQLSQEAPVYKRGKGRFRGRRLRLKDGHRISLPSPQHKITYQASPTLDKRSS-- 520
QY 392 MSEIEARIAATGNG--QPRRSIQDLVTGTPEGVSSRSSSPSVRMI----- 438
DB 521 ---DSGLCSPPGSPMLMLPLRLAIQ---LTPDENNKTRGRN-----WVFRQEDFEDVXR 567

QY 439 -----TTSGP-----TSEKP-----TRSHPWTPDDSTDTNGSDNSIPNAVLTLDH 478
DB 568 SFKKKGCTWGPSSVQTKERPEGRERVRPLSDGNSFW-----SLLIKSQKTTPLASLFVDQ 623
QY 479 QLOPLAPCPNKSMAVFE---CHCKMAQEMKQVETIALL-----QRQELVAE 526
DB 624 -----PSSCEQKLVPEGLEHRRKPKQTKPQQAHVGLPLCKDSQREDSSEAE 670

Search completed: December 9, 2002, 22:58:22
Job time : 67.5512 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:48:29 ; Search time 24.5928 Seconds
(without alignments)
692.718 Million cell updates/sec

Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTASAASSSSSSAGEMIE.....QCKQLEVRISQQKROGTS 579

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/prodata/1/iaa/5A COMB.pbp:**
- 2: /cgn2_6/prodata/1/iaa/5B COMB.pbp:**
- 3: /cgn2_6/prodata/1/iaa/6A COMB.pbp:**
- 4: /cgn2_6/prodata/1/iaa/6B COMB.pbp:**
- 5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pbp:**
- 6: /cgn2_6/prodata/1/iaa/backfiles1.pbp:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score		Query Match	Length	DB ID	Description
	%					
1	3014	100.0	579	4	US-09-529-279-4	Sequence 4, Appli
2	3014	100.0	590	4	US-09-529-279-15	Sequence 15, Appli
3	481	16.0	455	3	US-09-221-235-5	Sequence 5, Appli
4	481	16.0	455	3	US-09-221-928-5	Sequence 5, Appli
5	481	16.0	455	4	US-09-221-527-5	Sequence 5, Appli
6	481	16.0	455	4	US-09-221-236-5	Sequence 5, Appli
7	481	16.0	455	4	US-09-221-416-5	Sequence 5, Appli
8	481	16.0	455	4	US-09-221-245-5	Sequence 5, Appli
9	481	16.0	455	4	US-09-163-115-5	Sequence 5, Appli
10	481	16.0	455	4	US-09-221-528-5	Sequence 5, Appli
11	481	16.0	455	4	US-09-593-553-5	Sequence 5, Appli
12	481	16.0	455	4	US-09-221-237-5	Sequence 5, Appli
13	418	13.9	835	4	US-09-291-839-2	Sequence 2, Appli
14	412	13.7	668	1	US-08-205-018-2	Sequence 2, Appli
15	412	13.7	859	1	US-08-395-580-2	Sequence 2, Appli
16	412	13.7	859	5	PCT-US95-02792-2	Sequence 2, Appli
17	398	13.2	821	1	US-07-928-464-2	Sequence 2, Appli
18	398	13.2	821	1	US-08-003-311B-2	Sequence 2, Appli
19	398	13.2	821	1	US-08-261-432-2	Sequence 2, Appli
20	398	13.2	821	5	PCT-US93-07347-2	Sequence 2, Appli
21	393	13.0	263	3	US-09-035-706-5	Sequence 5, Appli
22	393	13.0	263	3	US-08-955-841-5	Sequence 5, Appli
23	393	13.0	263	4	US-09-390-425-5	Sequence 5, Appli
24	393	13.0	263	4	US-09-566-906-5	Sequence 5, Appli
25	390.5	13.0	787	4	US-09-188-930-334	Sequence 334, App
26	380.5	12.6	536	4	US-09-188-930-185	Sequence 185, App
27	379.5	12.6	1584	4	US-09-457-040B-27	Sequence 27, Appli

28	363.5	12.1	527	4	US-08-426-509A-10	Sequence 10, Appli
29	363.5	12.1	527	5	PCT-US95-05008-10	Sequence 10, Appli
30	358.5	11.9	620	4	US-08-426-509A-9	Sequence 9, Appli
31	358.5	11.9	620	5	PCT-US95-05008-9	Sequence 9, Appli
32	356.5	11.8	1052	3	US-08-863-118-2	Sequence 3, Appli
33	356	11.8	625	1	US-08-391-615-3	Sequence 1, Appli
34	355.5	11.8	983	3	US-08-167-919A-10	Sequence 10, Appli
35	355.5	11.8	983	3	US-08-715-106-10	Sequence 10, Appli
36	354.5	11.8	983	2	US-08-449-645A-21	Sequence 21, Appli
37	354.5	11.8	983	2	US-08-702-367A-21	Sequence 21, Appli
38	354.5	11.8	983	5	PCT-US95-04681-21	Sequence 21, Appli
39	351.5	11.7	269	2	US-07-857-224B-79	Sequence 79, Appli
40	349.5	11.6	304	2	US-08-701-191A-27	Sequence 27, Appli
41	349.5	11.6	1052	3	US-08-863-118-1	Sequence 1, Appli
42	349.5	11.6	1052	4	US-09-377-310-2	Sequence 2, Appli
43	346.5	11.5	982	2	US-08-673-789-4	Sequence 4, Appli
44	345.5	11.5	680	1	US-08-542-363-4	Sequence 4, Appli
45	345.5	11.5	680	4	US-09-100-089-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-529-279-4

Query Match		100.0%;	Score 3014;	DB 4;	Length 579;
Best Local Similarity		100.0%;	Pred. No. 3.8e-222;		
Matches 579;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSTASAASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKV	60		
Db	1	MSTASAASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKV	60		
Qy	61	AIKQIESESEKAFIVELRQLSRVNHNIKLYGACLNPCVCLMVEYAEGLSVNLHGAE	120		
Db	61	AIKQIESESEKAFIVELRQLSRVNHNIKLYGACLNPCVCLMVEYAEGLSVNLHGAE	120		
Qy	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGVLCIDFGTAC	180		
Db	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTGVLCIDFGTAC	180		
Qy	181	DIQTHMTNNKSAAMWAPPEVPEGSNYSEKCDVFSWGIIWEVITRRKPPFDEIGGPAPRIM	240		
Db	181	DIQTHMTNNKSAAMWAPPEVPEGSNYSEKCDVFSWGIIWEVITRRKPPFDEIGGPAPRIM	240		
Qy	241	WAVNGTRPPLIKNLPKPIESLMTRCWSKDPSPORMSEEVKIMTHLMRYFPGADEPLQY	300		
Db	241	WAVNGTRPPLIKNLPKPIESLMTRCWSKDPSPORMSEEVKIMTHLMRYFPGADEPLQY	300		
Qy	301	PCQYSDGQSNATSSTGSMFMDIASNTSNKSDTNEQVPATNDTIKLESKLLKNQAKQQ	360		

Db 301 PCQYSDGSGNSATSTGSPMDIASNTSNKSDTNMEQVPAITNDTIKRLSKLKNQAKQ 360
Qy 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTTGNGQPRRSIODLTVTG 420
Db 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTTGNGQPRRSIODLTVTG 420
Qy 421 TEPGVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
Db 421 TEPGVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
Qy 481 QPLAPCPNKSESAFAVEQHCXMAOEYMKVOTETALLQKQELVAELDDKQOQNTSRL 540
Db 481 QPLAPCPNKSESAFAVEQHCXMAOEYMKVOTETALLQKQELVAELDDKQOQNTSRL 540
Qy 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
Db 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
RESULT 2
US-09-529-279-15
; Sequence 15, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-15
Query Match 100.0%; Score 3014; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 3.9e-222;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 361 SESGRSLGASHGSSVSESLPPTSEGRKMSADMEIEARIAATTTGNGQPRRSIODLTVTG 420
Qy 421 TEPGVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
Db 421 TEPGVSSRSSSPSVMTTSGPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQL 480
Qy 481 QPLAPCPNKSESAFAVEQHCXMAOEYMKVOTETALLQKQELVAELDDKQOQNTSRL 540
Db 481 QPLAPCPNKSESAFAVEQHCXMAOEYMKVOTETALLQKQELVAELDDKQOQNTSRL 540
Qy 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
Db 541 VOEHKKLIDENKSLSTYYQCKKQLEVIRSQQKRGOTS 579
RESULT 3
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; EARLIER APPLICATION NUMBER: 1998-12-28
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-5
Query Match 16.0%; Score 481; DB 3; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;
Qy 27 NFEIDYKEIEVEBVGAGFGVVCXKAKW--RAKDAIK--QISESERKAFIVELROL 81
Db 7 SFVQKFDLDQFFENCGGSGFSGVYRAKMWISQKREAVAKLTKIEKEAF-----IL 57
Qy 82 SRVNHPIVVKLYGACLP--VCLVMEYAEGSLYINVHGA--AEPLPYTTAAHMSWCLQC 137
Db 58 SVLSHRNIIQFYGVILPEPNYGIYTEYASLSGLYDINSNRSEBM--DMDHIMTATDV 114
Qy 138 SOGVAYLHSMQKALIHRLDKPNNLLVAGTVLKICDFGTACDIQTHMTNNK--GSAAM 195
Db 115 AKGMHYILMEAPVAVYIHRDLKSRNVVIAADG-VLKICDFG-ASRFHHHTHMSLVGTFFW 172
Qy 196 MAPEVFECSNYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM-AYANGTRPPLIKN 254
Db 173 MAPEVIGSLPVSERCDITYSGVVLVMEMLTREVPFKGEG--LQVAVLVEKNERLTPSS 230
Qy 255 LKPIESLMTCSKSDSQSPSMEEIVKIMTHLMRYFGADEPLQYCYQSDGQNSAT 314
Db 231 CPRSFABELIHCWEADAKKRPSFKOIIISL-----ESMSNDT- 267
Qy 315 STGSPMDIASNTSNKSDTNMEQVPAITNDTIKRLSKLKNQAKQSESGRLSL--GAS 371
Db 268 ---SLPDKCNFLLHKKWEMRCE-IEATLERUKLEBLSFEQGLKERERLKNMEQKLT 323
Qy 372 HGSVSESLPPTS-----EGKMSADMEIEARIAAT-TGNGQPRRSIODLTVTG 421
Db 324 EGSNTPULLPLAAMSESESYESTBESNSAEMSCQITATSNGBGHWNSLOQMLMGF 383
Qy 422 EPGQVSSRSSSPSV 435
Db 384 --GDIFSMKAGAV 395
RESULT 4


```

Db      115 AKGHNYLHMEAPVAVIHRDLKSRNVVIAAG-VLIKIDFG-ASRFHNHTHMSLVTFPM 172
QY      196 MAPRVFGSNVSEKCDVFSWGIILMEVITRRKPEDEIGCAPRIMV-AVANGTRPPLIKN 254
Db      173 MAPVIGQLPVSEICDITYGVVIMEMLTREVPFKGLEG--LQVAMLVKERNRLTPSS 230
QY      255 LPKIEISLMTRCWSKDPSPQSPMEIVKIMTHLMRYFPADDEPLQYPCOYSDGQSNAT 314
Db      231 CPRFAELLHQCWADAKKRSPFOIISL-----ESMSNDT- 267
QY      315 STGSFMDLAINSTNKSOTNMEOVPATNDTKRLSKLLNQAQOSESRLSL---GAS 371
Db      268 ---SLPDKCNSFLHNKAEWRCE-IEATLERKLIERLRSKEQELKREERRLIMMGQKLT 323
QY      372 HGSSVESLPLPTS-----EGKRMGADMSLEIARIAAT-NGQPRRRSIODLVYTG 421
Db      324 EQSNTPLLLPILAAHMSSESYFESYTEESNSAEMSCQITATISNGEIGHMNPLOAMLMGF 383
QY      422 EPGGVSSRSSSPSV 435
Db      384 --GDIFSNMKAGAV 395

```

RESULT 7
US-09-221-416-5
Sequence 5, Application US/09221416

Patent No.6153417
GENERAL INFORMATION:
APPLICANT: Action, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,416
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ. ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

```

Query Match	16.0%	Score 481;	DB 4;	Length 455;
Best Local Similarity	30.4%;	Pred. No. 6-2e-29;		
Matches 132; Conservative	81;	Mismatches 151;	Indels 70;	Gaps 18

```

0Y 27 NEEEDYKIEIEVEEVGAGFVCCAKM--RAKDVAIK--QJESSEKFAIVELROL 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 SSVQJIKFDLDLOFEENCGGSGFSYIRAMNIOQDEKVAIVKLLIKTEKAE-----IL 57
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 82 SKVNHPIVILKYGLACLP--VCLVWEYAEGSLYNVLHG--AEFLPYITAAHMSCLQC 137
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SVLSHRNIIQFVGVLLEPPNNGIVTEYSLSLSDYDINSNSEEM--DMDHMTWADTV 114
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 138 SOGVAVLHSMQKALIHBDIKRPNLLVAGTVLKCIDCFACDIOQHTMTNNK--GSAM 195
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 AQGMVLIHAEAVKVIHRDLKSRNVIANADG-VLKICDPG-ASGFHHHTHMSLVGTFFM 172
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 196 MPEVEEGSNYSSEKDFEWSGCIILWEVITRRKPEDEIOGPAFRIMV-AVHNGTRPPLIKN 254
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 MAPEVIGSLPVSELTCDTYSYGVVLNEMLTREVPFGLEG--LQVAMVLVEKNERTLTPSS 230
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 255 LPKPIESLMTRCWSKDPORPSEMEIYKIMTHLNRKFFGADEPLQYFCQYSDSEQNSAT 314
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 CPRSFAELLHQCEWBAKRRPSFKOIIISIL-----ESMSNDT- 267
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 315 STGCSFMDIASITNMSKSDTNMEQVATNDTIKLESKLLKQAKQOSESGLSLP---GAS 371
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 268 ---SLPDCKNSFLNHAKEWRCF-IEATLERLKKLEPRDSFPEQELKERRRLKMWBOXL 323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
0Y 372 HGGSVSESLPPTS-----EGKMSADMEIEARIATTT-GNGQPRRNSIODLVYGT 421

```

Db 334 EOSNPILLPLFAANKSESYFESKTEBSNSHEMSCQITAINSGHGHNPSLQAMLMGF 382

Qy 422 EPGVSSRRSSPSV 435

Db 384 --GDIFSNNKAGAV 395

RESULT 8

US-09-221-245-5
; Sequence 5, Application US/09221245

Patent No. 6180358
; GENERAL INFORMATION:

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

```

; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245

```

```

; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115

```

```

; EARLIER FILING DATE: 1998-0
; NUMBER OF SEQ ID NOS: 15

```

; SOFTWARE: Pa
; SEQ ID NO 5

```

; LENGTH: 455
; TYPE: PRT

```

ORGANISM: Homo sapiens
US-09-221-245-5

US-09-221-245-5

Query Match	16.0%;	Score 481;	DB 4;	Length 455;
-------------	--------	------------	-------	-------------

Best Local Similarity 30.4%; Freq. NO. 0.26-23;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18

27 NEEIDYKEIEVEEVGGAFCVCKAKM--RAKDVAK--QIESESEKAFIVELROL 81

```

Db      7 SFVQIKRDDLQFFENCGGSGFSGVIYRAKMWISQDKREAVAKKLKLEKEAE-----IL 57

```

82 SRVNHPIVKLYGACLP--VCLVMEYAEGLYNVLHG--AEPLPYTAHAHMSWCLOC 13

58 SVLSHRNIQFYGVILLEPPNYGIVLEYASLSGLDYINSNRSEEM---DMDHITWATDV 11

138 SGGVAYLHMOPKALIRDLKPFNLLVAGTVLKICDFGACDIQTHMTNKK--GSAAW 19

115 AKGMHYLMEAPVKVIHRDLKSRNVIADG-VLKI CD FG-ASRFHNHTTHNSLVGTFPW 17

196 MAPEVEGSEKCDVFSWGLILMEVITRRKPEDEIGGAFRIMV-AVHNGTRPPLIKN 25

Db 173 MAPEVIQSLPVSETCTIYSYGVVLMEMLTREVPKGLGEG--LQVAMLVKEKNERLTIPSS 233

255 LPKPISLMTKCSKDBSQPSMEEYVKIMTHLNRFFPGADEPLQYPCQYSDGQNSAT 31

```

Db      231 C P R S F A E L L H Q C W E A D A K K R P S F K Q I I S I L ----- E S N S N D T - 26

```

315 STGSFMDIASTNTSNKBDPTNMEQVPATNDTIKRLESKLKNQAKQSESGRISL---GAS 37

Db 268 ---SLPDKNSFLHNKMRCE-IEATLERLKKLERDLSFKGEJLERERRLKWEQKLT 32

QY 372 HGSVESTLPPTS-----EGKRN\$ADMESEIARIAAT-GNGQPRRSIQDLVTGT 42

Db 324 EQSNTPLLLPLARMSSESYFESKFTESNSAEMSCQITATNSNGHGGMNPQLQAMLMGF 38

Qy 422 EPGVSSRSSPSV 435

Db 384 --GDIFSINKAGAV 395

DEBIT 0

ASU001 3
US-09-163-115-5
Sequence 5 Amplification ITS/09163115A

Patent No. 6183962

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREOF

FILE REFERENCE: MNI-050

```
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

Query Match          16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGRGAFGVVCKAKW--RAKDVAIK---QIESSEKAFIVELROL 81
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 SFVQIKFDDLOFFENCNGGSGFVSRAKWI SODKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHPIVVKLYGACLNLP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMSWCLQC 137
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASGLSLYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 173 MAPEVIQSLPVSETCDIYSYGVVLWEMLTREVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQYSDGQSNAT 314
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 CPRSFAELLHQWEADAKRPSFKQIISIL-----ESMSNDT- 267

QY 315 STGSFMDIASNTNSKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRLSL---GAS 371
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 268 ---SLPKCNKSNFLHNKAWRCE-TEATLERLKKLERDLSFKEQELKERERLKNWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 324 EQSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQITATSNEGHGWNPSLQAMLMGF 383

QY 422 EPGQVSSRSRSPSV 435
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 --GDIFSMNKAGAV 395

RESULT 11
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

Query Match          16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGRGAFGVVCKAKW--RAKDVAIK---QIESSEKAFIVELROL 81
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 7 SFVQIKFDDLOFFENCNGGSGFVSRAKWI SODKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHPIVVKLYGACLNLP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHAMSWCLQC 137
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASGLSLYDINSNRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFTACDIQTHMTNNK--GSAAW 195
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 115 AKGMHYLHMEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172

QY 196 MAPEVFEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 173 MAPEVIQSLPVSETCDIYSYGVVLWEMLTREVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPIESLMTRCWSKDPQRSMEIEIVKIMTHLMRYFFGADEPLOYPCQYSDGQSNAT 314
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 231 CPRSFAELLHQWEADAKRPSFKQIISIL-----ESMSNDT- 267

QY 315 STGSFMDIASNTNSKSDTNMEQVPATNDTIKRLSKLLKNQAKQOSESGRLSL---GAS 371
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 268 ---SLPKCNKSNFLHNKAWRCE-TEATLERLKKLERDLSFKEQELKERERLKNWEQKLT 323

QY 372 HGSSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRSIQDLTVTGT 421
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 324 EQSNTPLLLPLAARMSEESYFESKTEESNSAEMSCQITATSNEGHGWNPSLQAMLMGF 383

QY 422 EPGQVSSRSRSPSV 435
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 384 --GDIFSMNKAGAV 395

RESULT 10
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

Query Match          16.0%; Score 481; DB 4; Length 455;
Best Local Similarity 30.4%; Pred. No. 6.2e-29;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NFEIDYKEIEVEEVGRGAFGVVCKAKW--RAKDVAIK---QIESSEKAFIVELROL 81
DQ : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 27 NFEIDYKEIEVEEVGRGAFGVVCKAKW--RAKDVAIK---QIESSEKAFIVELROL 81
```

```

Db      221 CRRSEAEHLHQCEWADAKRPSFKQIISIL-----ESMSNDT- 267
QY      315 STGSGMDIASTNTSKSDPTNEQVPATNDTIKRLSKLLKNQAKQOESGRLSL---GAS 371
Db      268 ---SLPDKCNLSLHNAKWRCE-IEATLERLKKLERDLSFKEQELKEERLKKMEQKLT 323
QY      372 HGSSEVESLPTPTS-----EGKMSADMSSEIARIAATT-GNGQPRRSIOTLTVTGT 421
Db      324 EGSNTPLLPLLAARMSEESYFESKTESNSAEMSCQITATNSGEGHGMNPISLOAMLMGTF 383
QY      422 EPGQVSSRSSSPSV 435
Db      384 --GDIFSMNKAGAV 395

```

RESULT 12
US-09-221-237-5

```

; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

```

Query Match 16.0%; Score 481; DB 4; Length 455;

Best Local Similarity 30.4%; Pred. No. 6,2e-29; Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

```

QY      27 NFEIYKEIEVEEVGGAFGVVCRAKVM--RAKQVAIK---QIESESRKAFIVELROL 81
Db      7 SFOQIKFDDLDQFPENCNGGSGFSVYRAKMI SQKXEVAVKKLKIKEAL-----IL 57
QY      82 SRVNHFNIVLYGACLN--VCLVMEYABGSLYNVLHG--AEPLPYTAHAMSWCLQC 137
Db      58 SVLSHNITIQFYGVILLEPRYGI VTEYASLSGLDYINNSRDEM---DMDHIMTWDTV 114
QY      138 SOGVAIYHSMQKALIHRLDKPRLNLLVAGGYLTKICDFGTACDIQTHTNTNK--GSAAM 195
Db      115 AKGMHYLHMEAPVKVYIHRDLKSRNVVIAADG-VLKI CD FG-ASRFHNTHTHMSLVGTFFP 172
QY      196 MAPEVEGSSVYSKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIM-AVHNGTRPPLIKN 254
Db      173 MAPEVLIQSLVSETCTDTISYIGVILMELTREVFKGLEG--LOVALLVYKXERLITPS 230
QY      255 LKPRISLITRCWSKDPQSRPSMEIYKIMTHLMRYFPGADEPLYPCQYSDGGSNAT 314
Db      231 CRRSPAEHLHQCEWADAKRPSFKQIISIL-----ESMSNDT- 267
QY      315 STGSGMDIASTNTSKSDPTNEQVPATNDTIKRLSKLLKNQAKQOESGRLSL---GAS 371
Db      268 ---SLPDKCNLSLHNAKWRCE-IEATLERLKKLERDLSFKEQELKEERLKKMEQKLT 323
QY      372 HGSSEVESLPTPTS-----EGKMSADMSSEIARIAATT-GNGQPRRSIOTLTVTGT 421
Db      324 EGSNTPLLPLLAARMSEESYFESKTESNSAEMSCQITATNSGEGHGMNPISLOAMLMGTF 383
QY      422 EPGQVSSRSSSPSV 435
Db      384 --GDIFSMNKAGAV 395

```

RESULT 13

```

US-09-291-839-2
; Sequence 2, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-068
; CURRENT APPLICATION NUMBER: US/09/291,839A
; CURRENT FILING DATE: 1999-04-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-291-839-2

```

Query Match 13.9%; Score 418; DB 4; Length 835;
Best Local Similarity 31.6%; Pred. No. 9e-24;
Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;

```

QY      31 IDYKEIEVEEVGGAFGVVCRAKMPRAKVAIKQIE-----SESEKAFIVELROLSRV 85
Db      458 LQSEIEFHEHIIIGSGFGKVKGRCKNKI VAIKRYRANTYCSKSDVDFCREVSLICQLN 517
QY      86 HPIIVKLYGACLN--PVCVMEYABGSLYNVLHGAEPLPYTAHAMSWCLQCSQVA 142
Db      518 HPCVIOFGACLNDDPSQFALVTQYIGGSLFSLH--EQKRILDIQSKLIIADVAKGME 575
QY      143 YHSM-OPKALIHRLDKPRLNLLVAGGYLTKICDFGTACDIOT---HMTNKGSAAMNA 197
Db      576 YLHWLTOP--IHRDLNSHILLYEDGHAV-VADFGESRLOSLEDNMTKQGNLRMA 632
QY      198 PEVF-EGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNLP 256
Db      633 PEVFTQCTRTYIKADVFSYALCLMEILITGEIPRAHLKPAADAAVMYHH-IRPPIGYSIP 691
QY      257 KPISLITRCWSKDPQSRPSMEIYKIMTHLMRYFPGADPEPLYPCQYSDGGSNATST 316
Db      692 KPISLIRGNMACPEGRPFSEVVMKLECL-----CNIELMSFA-----SSNS 737
QY      317 GSFMDIASTN-TSKSDPTNEQVPATNDTIKRLSKLLKNQAKQOESGRLSGASHGSS 375
Db      738 GSLSPSSSSDCLVNRGGPGRSHVALLRSPELEYALNARSYALSSQAGYS---SGGLS 794
QY      376 VESL 379
Db      795 LEEM 798

```

```

RESULT 14
US-08-205-018-2
; Sequence 2, Application US/08205018
; Patent No. 5554523
; GENERAL INFORMATION:
; APPLICANT: Reddy, Usharani R.
; TITLE OF INVENTION: No. 5554523el Protein Kinase, Nucleic Acid
; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
; ADDRESS: No. 5554523is
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

```



```
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/205,018
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Gaumont, Rebecca R.
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 668 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-205-018-2

Query Match 13.7%; Score 412; DB 1; Length 668;
Best Local Similarity 24.6%; Pred. No. 1.9e-23;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEVGRGAFVGVCKAKWRAKDVAIKQIESESEKAFIVELRQLSRVNHNI 89
DB 119 EYPPEILDQWVGSGAQAQVFLGRFHGEEVAVKVRDLKE-----TDIKHLRKLKHPNI 173
QY 90 VKLYGACINP--VCLVMEYAEAGSLYLVHGAEPPLPYTTAAHAMSWCLQCSQGVAYLHSM 147
DB 174 ITFKGVCTQAPCYCILMEFCAQQQLYEVLRAQRPV---TPSLLDVMSMGIAAGMNYLHLH 230
QY 148 QPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAMMAPEVEFGSN 205
DB 231 K---IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVLLWELLTGEIPYKDVDSA--IIVGVGNSLHLPVPSSCPDGFKILLR 344
QY 265 RCWSKDFSRPSMEEIVKIMTHLMRYFPGADEPLOYPCQYSDGQSNSTATSGSFMDIAS 324
DB 345 QCWNSKPNRPSFRQ---ILLHL-----DIAS 368
QY 325 TNT-SNKSdT-----NMEQVPATNDTIKRLESKLL----- 353
DB 369 ADVLSTPQETVFKSQAEWREVKLHFKEIKSEGTCLHRLHEBELVMRRRELRHALDIREH 428
QY 354 -----PPTSEGRMSADMEIEARI-----AATTGQPPRRRSIQDLTVTGTPEQVSS 428
DB 489 LIKKRNPQNLSPHSRQPDILKAESLLPKLDAALSGVGLP-----GCPKAPPSGR--S 540
QY 429 RSSSPSVRMITTSPTSEKP---TRSHPTPDDSDTDNG-----SDNSIPMAVLTLDHQ 479
DB 541 RRGTRHRKASAKSCGDLPLRLTAVPPHFGPGSGGLGGPSAWACPPALRGLLHD 600
QY 480 L 480
DB 601 L 601

RESULT 15
US-08-395-580-2
; Sequence 2, Application US/08395580
; Patent No. 5676945
; GENERAL INFORMATION:
; APPLICANT: Usharani R. Reddy, David Pleasure and the Children's
; APPLICANT: Hospital of Philadelphia
; TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid

; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESS: Woodcock Washburn Kurtz Mackiewicz and No. 5676945rls
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 Kb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/395,580
; FILING DATE: herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/205,018
; FILING DATE: 01-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca L. Ralph (formerly Gaumont)
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-395-580-2

Query Match 13.7%; Score 412; DB 1; Length 859;
Best Local Similarity 24.6%; Pred. No. 2.7e-23;
Matches 133; Conservative 79; Mismatches 181; Indels 148; Gaps 20;

QY 30 EIDYKEIEVEVGRGAFVGVCKAKWRAKDVAIKQIESESEKAFIVELRQLSRVNHNI 89
DB 119 EYPPEILDQWVGSGAQAQVFLGRFHGEEVAVKVRDLKE-----TDIKHLRKLKHPNI 173
QY 90 VKLYGACINP--VCLVMEYAEAGSLYLVHGAEPPLPYTTAAHAMSWCLQCSQGVAYLHSM 147
DB 174 ITFKGVCTQAPCYCILMEFCAQQQLYEVLRAQRPV---TPSLLDVMSMGIAAGMNYLHLH 230
QY 148 QPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAMMAPEVEFGSN 205
DB 231 K---IIHRDLKSPN-MLITYDDVVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
QY 206 YSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVLLWELLTGEIPYKDVDSA--IIVGVGNSLHLPVPSSCPDGFKILLR 344
QY 265 RCWSKDFSRPSMEEIVKIMTHLMRYFPGADEPLOYPCQYSDGQSNSTATSGSFMDIAS 324
DB 345 QCWNSKPNRPSFRQ---ILLHL-----DIAS 368
QY 325 TNT-SNKSdT-----NMEQVPATNDTIKRLESKLL----- 353
DB 369 ADVLSTPQETVFKSQAEWREVKLHFKEIKSEGTCLHRLHEBELVMRRRELRHALDIREH 428
QY 354 -----PPTSEGRMSADMEIEARI-----AATTGQPPRRRSIQDLTVTGTPEQVSS 428
DB 489 LIKKRNPQNLSPHSRQPDILKAESLLPKLDAALSGVGLP-----GCPKAPPSGR--S 540
QY 429 RSSSPSVRMITTSPTSEKP---TRSHPTPDDSDTDNG-----SDNSIPMAVLTLDHQ 479
```

Db 541 RRGKTRHRKASAKSGCDLPGLRTAVPHEPGSGSPGILGSGPSAMEACPPALRGGLHD 600
Qy 480 L 480
Db 601 L 601

Search completed: December 9, 2002, 23:00:36
Job time : 28.5928 secs

GenCore version 5.1.1.3
 Copyright (c) 1993 - 2002 Compugen Ltd.
 OM protein - protein search, using sw model
 Run on: December 9, 2002, 22:53:24 ; Search time 116.014 Seconds
 (without alignments)
 81.062 Million cell updates/sec

Title: US-09-830-144-2
 Perfect score: 3014
 Sequence: 1 MSTAASASSSSSSAGEMIE.....CKKQLEIVRSQQKRGQTS 579

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000
 Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Published Applications AA:*

```

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/ECT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/ECTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US03_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	579	US-10-158-895-4	Sequence 4, Appli
2	3014	100.0	590	US-10-158-895-15	Sequence 15, Appli
3	481	16.0	455	US-09-757-982-5	Sequence 5, Appli
4	461.5	15.3	1036	US-10-014-882-2	Sequence 2, Appli
5	418.5	13.9	394	US-09-862-027-19	Sequence 19, Appli
6	418	13.9	835	US-09-947-199-2	Sequence 2, Appli
7	416	13.8	328	US-03-862-027-18	Sequence 18, Appli
8	414	13.7	835	US-09-947-199-8	Sequence 8, Appli
9	408.5	13.6	966	US-09-771-161A-197	Sequence 197, App
10	393	13.0	263	US-09-840-704-5	Sequence 5, Appli
11	392	13.0	850	US-09-904-389-2	Sequence 2, Appli
12	369	12.2	425	US-03-828-313-29	Sequence 29, Appli
13	363.5	12.1	527	US-09-977-269-10	Sequence 10, Appli
14	360.5	12.0	277	US-09-882-166-4	Sequence 4, Appli
15	359	11.9	278	US-09-842-582-4	Sequence 4, Appli
16	359	11.9	278	US-09-797-039-13	Sequence 13, Appli
17	359	11.9	278	US-09-922-138-18	Sequence 18, Appli
18	359	11.9	278	US-09-922-138-27	Sequence 27, Appli
19	359	11.9	278	US-09-910-150-17	Sequence 17, Appli

20	359	11.9	278	10	US-09-910-150-31	Sequence 31, Appli
21	358.5	11.9	277	10	US-09-815-915-13	Sequence 13, Appli
22	356.5	11.9	620	10	US-09-977-269-9	Sequence 9, Appli
23	355	11.8	278	10	US-09-815-915-16	Sequence 16, Appli
24	354.5	11.8	983	10	US-09-771-161A-227	Sequence 227, App
25	351	11.6	273	10	US-09-922-138-11	Sequence 11, Appli
26	350	11.6	271	10	US-09-799-875-28	Sequence 28, Appli
27	350	11.6	272	10	US-09-780-949-4	Sequence 4, Appli
28	350	11.6	272	10	US-09-910-150-27	Sequence 27, Appli
29	349.5	11.6	1052	10	US-09-757-100B-2	Sequence 23, Appli
30	348.5	11.6	279	10	US-09-799-875-23	Sequence 2, Appli
31	345.5	11.5	680	10	US-09-827-949-4	Sequence 4, Appli
32	345.5	11.5	1620	10	US-09-827-949-2	Sequence 2, Appli
33	343	11.4	272	10	US-09-815-915-14	Sequence 14, Appli
34	341	11.3	1276	10	US-09-982-610-24	Sequence 24, Appli
35	340	11.3	628	10	US-09-862-027-48	Sequence 48, Appli
36	338.5	11.2	764	10	US-09-925-302-714	Sequence 714, App
37	337.5	11.2	928	8	US-08-578-684-2	Sequence 2, Appli
38	335.5	11.1	507	10	US-09-977-269-2	Sequence 2, Appli
39	333	11.0	1308	10	US-09-940-101-2	Sequence 15, Appli
40	331.5	11.0	280	10	US-09-515-806-15	Sequence 36, Appli
41	329.5	10.9	1104	10	US-09-982-610-36	Sequence 10, Appli
42	329	10.9	265	10	US-09-797-039-10	Sequence 3, Appli
43	329	10.9	675	9	US-10-186-399-3	Sequence 4, Appli
44	329	10.9	675	10	US-09-977-269-4	Sequence 6, Appli
45	323.5	10.7	271	10	US-09-840-704-6	

ALIGNMENTS

RESULT 1
 US-10-158-895-4
 ; Sequence 4, Application US/10158895
 ; Patent No. US20020155624A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ONO, KOICHIRO
 ; APPLICANT: TSUCHIYA, MASAYUKI
 ; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
 ; FILE REFERENCE: 053466/0278
 ; CURRENT APPLICATION NUMBER: US/10/158,895
 ; CURRENT FILING DATE: 2002-06-03
 ; PRIOR APPLICATION NUMBER: US/09/529,279
 ; PRIOR FILING DATE: 2000-04-11
 ; PRIOR APPLICATION NUMBER: PCT/JP98/04796
 ; PRIOR FILING DATE: 1998-10-22
 ; PRIOR APPLICATION NUMBER: JP 9/290188
 ; PRIOR FILING DATE: 1997-10-22
 ; NUMBER OF SEQ ID NOS: 48
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 579
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-158-895-4

Query Match 100.0%; Score 3014; DB 9; Length 579;
 Best Local Similarity 100.0%; Pred. No. 8,1e-176;
 Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSTAASASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKRAKDV	60
Db	1	MSTAASASSSSSSAGMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKRAKDV	60
QY	61	AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNIPVCLVMEYAEAGSLYNVLHGA	120
Db	61	AIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNIPVCLVMEYAEAGSLYNVLHGA	120
QY	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALITHRLKPPNLLLVAGGTVLKICDFGTAC	180
Db	121	PLPYTTAAHMSWCLQCSQGVAYLHSMQPKALITHRLKPPNLLLVAGGTVLKICDFGTAC	180

QY 181 DIQHTMTNNKGSAAWMAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAWMAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPFGADEPLQY 300
QY 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIATGNGQPRRSIIDLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIATGNGQPRRSIIDLTVTG 420
QY 421 TEPGQVSSSSSPSVNMTTSGPTSEKPRSHPMTPDDSDTNGSDNSIPMAVLTLDHQL 480
DB 421 TEPGQVSSSSSPSVNMTTSGPTSEKPRSHPMTPDDSDTNGSDNSIPMAVLTLDHQL 480
QY 481 QPLAPCPNKSMAVFEQHCMAQOYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
DB 481 QPLAPCPNKSMAVFEQHCMAQOYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
QY 541 VOEHKKULDENKSLSTYYQCKKQLEVIRSOQKROGTS 579
DB 541 VOEHKKULDENKSLSTYYQCKKQLEVIRSOQKROGTS 579

RESULT 2
US-10-158-895-15
Sequence 15, Application US/10158895
Patent No. US20020155624A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158, 895
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529, 279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-158-895-15

Query Match 100.0%; Score 3014; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 8.3e-176;
Matches 579; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGMTEAPSOVNFEEIDKEIEVEEVGRGAFGVYCKAKMRAKDY 60
DB 1 MSTASASSSSSSAGMTEAPSOVNFEEIDKEIEVEEVGRGAFGVYCKAKMRAKDY 60
QY 61 AIKQIESESERKAFIYELRQLSRNHNHIVKLYGACINPCLVMEYAEAGSLYNVLHGAE 120
DB 61 AIKQIESESERKAFIYELRQLSRNHNHIVKLYGACINPCLVMEYAEAGSLYNVLHGAE 120
QY 121 PLPYTTAAHMSWCLQSGGVAAYLHSMQPKALIHRLDKPNNLLVAGGYLTKICDPGTAC 180
DB 121 PLPYTTAAHMSWCLQSGGVAAYLHSMQPKALIHRLDKPNNLLVAGGYLTKICDPGTAC 180
QY 181 DIQHTMTNNKGSAAWMAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240

DB 181 DIQHTMTNNKGSAAWMAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPFGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPFGADEPLQY 300
QY 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
DB 301 PCQYSDGQSNASTSGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQ 360
QY 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIATGNGQPRRSIIDLTVTG 420
DB 361 SESGRSLGASHGSSVESLPPTSEGRKMSADMEIEARIATGNGQPRRSIIDLTVTG 420
QY 421 TEPGQVSSSSSPSVNMTTSGPTSEKPRSHPMTPDDSDTNGSDNSIPMAVLTLDHQL 480
DB 421 TEPGQVSSSSSPSVNMTTSGPTSEKPRSHPMTPDDSDTNGSDNSIPMAVLTLDHQL 480
QY 481 QPLAPCPNKSMAVFEQHCMAQOYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
DB 481 QPLAPCPNKSMAVFEQHCMAQOYMKVQTEIALLQKQELVABLDQDEKQDQNTSRL 540
QY 541 VOEHKKULDENKSLSTYYQCKKQLEVIRSOQKROGTS 579
DB 541 VOEHKKULDENKSLSTYYQCKKQLEVIRSOQKROGTS 579

RESULT 3
US-09-757-982-5
Sequence 5, Application US/09757982
Patent No. US20020094559A1
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/757, 982
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 09/163, 115
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRT
ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 16.0%; Score 481; DB 10; Length 455;
Best Local Similarity 30.4%; Pred. No. 2.3e-22;
Matches 132; Conservative 81; Mismatches 151; Indels 70; Gaps 18;

QY 27 NPEIDYKEIEVEEVGRGAFGVYCKAKM--RAKDAVK--QIESESEKAFIVELRQ 81
DB 7 SFVQIKFDDIQFFENCGGSGFYRAKWISODEKAVVKKLKEKAE-----IL 57
QY 82 SRVHNPNIVKLYGACINP--VCLVMEYAEAGSLYNVLHG--ABPLPYTTAAHMSWCLQ 137
DB 58 SVLSHRNLIQYGYILEBPNYGYTEYASLQSLYDYNRSRSEM--DMDHMTATDV 114
QY 138 SQGVAYLHSMQPKALIHRLDKPNNLLVAGGYLTKICDPGTACDIQHTMTNNK--GSAAW 195
DB 115 AKGMHVLHMEAPVAVYIHRDLKSRNVVIAADG-VLKIDFG-ASRFNHHTTHMSLVGTFPW 172
QY 196 MAPEVEFGSNYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIM-AVHNGTRPPLIKY 254
DB 173 MAPEVIGSLPYSEICDITYSYGVVLEMLTREVPFKGLEG--LQVAMLVVEKNEHRTIPSS 230
QY 255 LKPIESIMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPFGADEPLQYPCQYSDGQSNAST 314
DB 231 CFSRAELHQCWEADAKRPSFKQIISIL-----ESMSNDT- 267
QY 315 STGSFMDIASTNTSNKSDPTNMEQVATNDTIKRLSKLLKNQAKQSDGRLSL--GAS 371

```
Db 268 ---SLPDKCNFLHNKAERCE-IEATLERLKKLERDLSPKEQLKERERRLLKMWQKLT 323
QY 372 HGSVESLPPTS-----EGKMSADMSEIEARIAATT-GNGQPRRRRSIQDLTVTGT 421
Db 324 EQSNTPLLLPLAARMBESYFESKTEESNSAEMSCQITATSNGEGHGMNPSLOAMWLMGF 383
QY 422 EPQGVSRSSSPSV 435
Db 384 --GDIFSMNKAGAV 395

RESULT 4
US-10-014-882-2
; Sequence 2, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kiehe, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1el Human Kinase and Polynucleotides Encoding
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1036
; TYPE: PRT
; ORGANISM: homo sapiens
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(1036)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-10-014-882-2

Query Match 15.3%; Score 461.5; DB 12; Length 1036;
Best Local Similarity 26.7%; Pred. No. 9e-21;
Matches 171; Conservative 97; Mismatches 254; Indels 119; Gaps 26;

QY 22 PSQVLNFEIDYKEIEVEVRVGRGAFVVCCKAKRAKQVAIKQIESERKAFIV----- 76
Db 110 PRPSPVHVAFERLEKELIGAGFGQYRATWQGEVAVKAAQDPEDQAAAAAASVR 169
QY 77 -ELRQLSRVNHPIVKLYGACLNLP--VCLVMEYARGGSLYNVLHGAEPLPYVTA----- 127
Db 170 REARLFAMLRHENIETLRGVCLQQPHCLVLBFARGGALNRALANAAPDPRAPCPRA 229
QY 128 ----AHAM-SWCLQSQGVAYLHMQPKALIHRLDKPNLLLVAG-----GTVLKICD 175
Db 230 RRIPPHVLNVAVQIARGMLYLHHEAFVPIHLRLDKSSNILLLEKIEHDDICNKLKID 289
QY 176 FGTACD-IQHTMTNKGSAWMAPEVEGGSYSEKDFVSMGIIILWEVITRKPFDG 234
Db 290 FGLAREWHRTTKMSTAGTYAMWAPVIVKSSLSFGSDIWSGVLLWELLTGEVYRGIDG 349
QY 235 PAFRIMWAVHNGTRPPLIKNLKPISLMTRCWSDKPSQRPDS---MEEIVKIMTLMRY 290
Db 350 LAVAYGAVNKLTL-PIPTSCPEPAKLMKECWQDPIRPSFALLILEQLTAECAVWTE 408
QY 291 FPCA-----DELOYPCQYSDGQSNSATSGTFMFIADTNTSNKSDTNM----EQVP 339
Db 409 MPQESFHSMDDWKLEIQ-QMFDLRTKEKELRSREBELTRAALQKQSQEBLLKRREQQL 467
QY 340 ATN--DTIKRLSKLLKNAKQOQSSGRLSLGASHGSSVESLPTSEKRNMSADMSIEA 397
Db 468 AEREIDVLER-ELNLIIFQLNQEKPKVKRKKGKFKRSLLK----LKQHRISLP-SDFOH 521
QY 398 RIAATGNGQPRRRSIQDLTVTGTGPGQVSSRSSSPSVRMITTSQPTSEK----- 447
Db 522 KITVQASPNLDKRRSL-----NSSSSSPSPSPTMPLRAIQLTSDSNKTKWGRNTVFRQ 576
```

```
QY 448 -----PTRGHPWTPD-----DSTDTNG-----SDKSIPMAYLTLDHO-LOPLA-- 484
Db 577 EEFEDVKRNFKKGCTWGPNSIQMKDRITDCKERTRPISLDGNSPWSSTIILKNQKTMPLASL 636
QY 485 -----PCPNKESMAVPEQHCCKMAQYMKVQTEIALLL-----QRKQELVAELDQDEKQ 534
Db 637 FVDQPGSCSEPKLSPDGL-HRKPKQIKLPSQAYIDLPLGKDAQRENPAEAE-SWEEAAS 694
QY 535 QNTSRLVQEHKKLLDENKSLSTYYQCKKQLLEVIRSOQQR 575
Db 695 ANAATVSTIE---MTPTNSLS-----RSPQRKK 718

RESULT 5
US-09-862-027-19
; Sequence 19, Application US/09862027
; Patent No. US20020142428A1
; GENERAL INFORMATION:
; APPLICANT: Hodge, Martin R.
; TITLE OF INVENTION: No. US20020142428A1el Kinases and Uses Thereof
; FILE REFERENCE: 35800/234862
; CURRENT APPLICATION NUMBER: US/09/862,027
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 09/345,473
; PRIOR FILING DATE: 1999-06-30
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19
; LENGTH: 394
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-862-027-19

Query Match 13.9%; Score 418.5; DB 10; Length 394;
Best Local Similarity 36.6%; Pred. No. 1.2e-18;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

QY 35 EIEVEYVVGAFGVVCCKAKRAKQVAIK--QJESERKAFIVELRQ-----LSRVNHPN 88
Db 2 ELTLEEIIIGGFGKVYRATWIGDEVAVKAAHDPDEDISQTIENVRQEAFLFAMLKHPN 61
QY 89 IVKLYGACLNLP--VCLVMEYARGGSLYNVLHGAEPLPYVTAHAHMSWCLQSQGVAYLHS 146
Db 62 IIALRGVCLKEPNLCLVMEFARGGPNRVLSGKRIPPDI-----LVNWAYIARGMNYLHD 117
QY 147 MQPKALIHRLDKPNLLLV-----AGTVLKICDFGTACD-IQHTMTNKGSAWMAAP 198
Db 118 EAIVPIIHRDLKSSNILLIQKVENGDLSNKLKITDFGLAREWHRTTKMSAAGTYAMAP 177
QY 199 EVPEGSNYSKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPPLIKNLKP 257
Db 178 EVIRASWFSKGSVWSYGVLLWELLTGEVYRFGIDG--LRVAYGVAMNKLALPIPTCPE 235
QY 258 PIESLMTRCWSDKPSQRPDSMEIEIVKIMT 285
Db 236 PFAKLMECDWNPDPHGRSPSTNILDQLT 263

RESULT 6
US-09-947-199-2
; Sequence 2, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
```

PRIOR APPLICATION NUMBER: 09/458,457
 PRIOR FILING DATE: 1999-12-10
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 2
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-09-947-199-2

Query Match 13.9%; Score 418; DB 10; Length 835;
 Best Local Similarity 31.6%; Pred. No. 3e-18;
 Matches 115; Conservative 67; Mismatches 144; Indels 38; Gaps 13;

QY 31 IDYKEIEVEEVGRGAFGVYCKAKWRADVAIKOIE-----SESEKKAFFIVELRQLSRVN 85
 DB 458 IQLSEIEFEHIIIGSGSFGKYKGRCKNKIVALIKRYRANTYCKSDVDFCREVSTLCQLN 517
 QY 86 HPIVLYGACLN---PVCLVMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCLQCSQVA 142
 DB 518 HCVIQFVACLNDSQFAIVTQYISGSLFSLH--EQKRILDLQSKLIIVADVAKGME 575
 QY 143 YLHSM-QPPALIHRLDKPPNLLVAGGYVLCIDFGTACDIQT---HMTNKGSAAMA 197
 DB 576 YLHNLTOP--IHRDLNSHNLILYEDGHAV-VADFGESRFLQSLDEDNNTQOPGMLRWMA 632
 QY 198 PEYF-BGSNYSKCDVFSMGILLMEYITRRKPFDEIGCFAPRIMAVHNGTRPPLIKNLP 256
 DB 633 PEYFQCTYTTKADVFSTALCLMEILTEIPFAHKPAAAADAAVYH-IRPPIGYSIP 691
 QY 257 KPISLMTRCWSKDPQSPSMETVYKIMTHLMRYFPGADEPLQYPCQYSDEQSNSATST 316
 DB 692 KPISLLIRGMWACPEGRPEFSVVKLECL-----CNIELMSPA-----SSNS 737
 QY 317 GSGMDIASN-TSNKSDTVMEOVPATNDTIKRLSKLKNQAKQSESGSLGASHSS 375
 DB 738 GSLSPSSSDCLVNRGGRSHVALLRSRPELEYALNARSYALASQASQYS---SQGLS 794
 QY 376 VESL 379
 DB 795 LEBM 798

RESULT 7
 US-09-862-027-18
 Sequence 18, Application US/09862027
 Patent No. US20020142428A1
 GENERAL INFORMATION:
 APPLICANT: Hodge, Martin R.
 TITLE OF INVENTION: NO. US20020142428A1 Kinases and Uses Thereof
 FILE REFERENCE: 35800/234862
 CURRENT APPLICATION NUMBER: US/09/862,027
 CURRENT FILING DATE: 2001-05-21
 PRIOR APPLICATION NUMBER: US 09/345,473
 PRIOR FILING DATE: 1999-06-30
 NUMBER OF SEQ ID NOS: 82
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 18
 LENGTH: 328
 TYPE: PRT
 ORGANISM: C. elegans
 US-09-862-027-18

Query Match 13.8%; Score 416; DB 10; Length 328;
 Best Local Similarity 33.4%; Pred. No. 1.4e-18;
 Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;
 QY 1 MSTASAASSSSSAGMIEAPSQVLFEEIDYKEIEVEEVGRGAFGVYCKAKWRADY 60
 DB 22 MSTTSNESTSSSSNNS-----DQRLVLPDIQRDDIQVGDHIGVTFGAVFSGNWTLPG 76
 QY 61 AIKOIESESEKKAFFIVLRQ--LSRVNHPIVLYGACL---NPVCLVMEYAEAGSLYV 115

DB 77 SORTI---ALKVFLVEKEAEILSKIRKNIIOFYGICKATNDPFITYEAKSGSLYDF 133
 QY 116 LHGAELPYTTAAHMS-----WCLQSQGVAYLHMQPKALIHRLDKPPNLLVAGGT 169
 DB 134 IHSEESQSFASSSGNSFVUVKMASQIASGIQYLHYDAVDTIHRDLKSKNVVL-DKRL 192
 QY 170 VLKICDFGTACDIQTHMTNKN---GSAAMAPE-VFESSNYSKCDVFSMGILLMEVITR 225
 DB 193 VCKICDFGTSKDL-THSCAPSWGTAAMSPSEMILQSEGLTTATDVMSYGVVLEILSK 251
 QY 226 RKPDEIGCFAPRIMAV-HNGTRPPLIKNLPKPIESLMTRCWSKDPQSPSMETI 280
 DB 252 EYVYKQYS--EPIIFMTIQSGITTLALIPSCAPLKLQMSNCWMTPKDRAWRQI 305

RESULT 8
 US-09-947-199-8
 Sequence 8, Application US/09947199
 Patent No. US20020127684A1
 GENERAL INFORMATION:
 APPLICANT: Raju, Jeyaseelan
 TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
 FILE REFERENCE: MNI-06RCP2
 CURRENT APPLICATION NUMBER: US/09/947,199
 CURRENT FILING DATE: 2001-09-05
 PRIOR APPLICATION NUMBER: 60/111,938
 PRIOR FILING DATE: 1998-12-11
 PRIOR APPLICATION NUMBER: 09/291,839
 PRIOR FILING DATE: 1999-04-14
 PRIOR APPLICATION NUMBER: 09/458,457
 PRIOR FILING DATE: 1999-12-10
 NUMBER OF SEQ ID NOS: 9
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 8
 LENGTH: 835
 TYPE: PRT
 ORGANISM: Rattus norvegicus
 US-09-947-199-8

Query Match 13.7%; Score 414; DB 10; Length 835;
 Best Local Similarity 31.7%; Pred. No. 5.3e-18;
 Matches 120; Conservative 67; Mismatches 143; Indels 48; Gaps 15;

QY 20 EAPSQVLFEEIDYKEIEVEEVGRGAFGVYCKAKWRADVAIKOIE-----SESEKKA 74
 DB 451 ELPSRF---HLQSEIEFEHIIIGSGSFGKYKGRCKNKIVALIKRYRANTYCKSDVDF 506
 QY 75 IVELRQLSRVNHPIVLYGACL---NPVCLVMEYAEAGSLYVNLHGAEPLPYTTAAHAM 131
 DB 507 CREVSILOQLNHPCVVQPVGACLDDBSQFAIVTQYISGSLFSLH--EQKRILDLQSKL 564
 QY 132 SWCLQSQGVAYLHSM-QPPALIHRLDKPPNLLVAGGYVLCIDFGTACDIQT---HM 186
 DB 565 IIVADVAKGMEYHSLQP--IHRDLNSHNLILYEDGHAV-VADFGESRFLQSLDEDNM 621
 QY 187 TNNKGAAMAAPEYF-BGSNYSKCDVFSMGILLMEYITRRKPFDEIGCFAPRIMAVHN 245
 DB 622 TKQPGNLRWMAPEVFTQCTYTTKADVFYSYSLCLMEILTEIPFAHKPAAAADAAVYH 681
 QY 246 GTRPPLIKNLPKPIESLMTRCWSKDPQSPSMETVYKIMTHLMRYFPGADEPLQYPCQYS 305
 DB 682 -IRPPIGYSIPKPISSILIRGMWACPEGRPEFSVVKLECL-----CNVELMSPA--- 732
 QY 306 DEQGSNSATSTGSPMDIASNTSNKSDTVMEOVPATNDTIKRLSKL-LKNQAKQSESG 364
 DB 733 -----SGNSGSL-----SPSSSDCLSRGGRSHVALLRSRPELEYALNARSYAG 780
 QY 365 RSLGLASH---GSSVESL 379
 DB 781 WSQSVGTHSNPGLSLEEM 798

Best Local Similarity 32.1%; Pred. No. 1,2e-16;
Matches 93; Conservative 66; Mismatches 111; Indels 20; Gaps 12;

QY 17 ENIEAPSOV-LNFEF--IDYKEIEVEEVGRAPGVCKAKKRAADVAKIQIESSE---E 70
 Db 555 QLIKRPNELSLGLEDLVLPMTDLIREKIGAGSFGTVRGEMHSGDVAVKLTLEDQDFHPE 614
 QY 71 R-KAFIVELRQLSRVNHPNIVKLYGACLN--VCLVMEYAEAGSLYNYLH--GAEPLEPY 125
 Db 615 RNEFLREVAIKSLRHPNIVFPMGAVTKPPLSLVTEYLSNGSLYRLHKSQVMDI--- 671
 QY 126 TAAHMSWCLQSGQVAVYHSMQPKALIHRLDKPNNLLVAGCTVLYKICDQGT-TCDDIQT 184
 Db 672 DETRRINMAFDVAKGWNVYHRRDP--PIYHRDLKSPNLLVDKXYTV--KYCDFGLSLKART 729
 QY 185 HMTNKK--GSAAMMAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPPDEIGGAPARIMWA 242
 Db 730 FLSSKSAAGTPPEWMAPEVLRDEPSNEKSDVYSGVILMELATLLOQPCWNL--NPAQVVAAY 788
 QY 243 VHNQTRPPLIKNLPRPESLMTKRCWSKDPQSPRSMEEIVKIMTHLMRYFP 292
 Db 789 GFRGKRKLIDIPROVNPKLALIVACMADEPMKRPSSISIMETLKPMTKQAP 838

RESULT 12
 US-09-828-313-29
 Sequence 29, Application US/09828313
 Patent No. US20020059662A1

GENERAL INFORMATION:
 APPLICANT: COSTA e SILVA, OSMALDO DA
 APPLICANT: BOHNER, HANS J.
 APPLICANT: THIELEN, NOCHA VAN
 APPLICANT: CHEN, ROUYING
 APPLICANT: SARRIA-MILLAN, RODRIGO
 TITLE OF INVENTION: PROTEIN KINASE STRESS-RELATED PROTEINS AND METHODS OF
 TITLE OF INVENTION: USE IN PLANTS
 FILE REFERENCE: 16313-0032
 CURRENT APPLICATION NUMBER: US/09/828,313
 CURRENT FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/196,001
 PRIOR FILING DATE: 2000-04-07
 NUMBER OF SEQ ID NOS: 128
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 425
 LENGTH: 425
 TYPE: PRT
 ORGANISM: Physcomitrella patens
 US-09-828-313-29

Query Match 12.2%; Score 369; DB 10; Length 425;
 Best Local Similarity 33.8%; Pred. No. 1.3e-15;
 Matches 94; Conservative 51; Mismatches 101; Indels 32; Gaps 10;

QY 21 ABSQVLFNEFDYKEIEVEEVGRGAFGVCKAKKRAKDAVAKIQIESSEERAKFIY---- 76
 Db 131 APSEI-----ELDTSELIGKAGFETKALMGTPVAVKTIIPSLSNDRMVTKDQ 181
 QY 77 -ELRQLSRVNHPNIVKLYGACL--NPVCLVMEYAEAGSLYNYLHGAEPPLPYTTAAHMSM 133
 Db 182 HEVOQLVKVHPNIVQFLGAVTRQRPMLVTEFLAGSDHLQLRSN---PNLAPDRIYKY 238
 QY 134 CLOCSQGVAVYHSMQPKALIHRLDKPNNLLVAGTVLYKICDQGT---CDIQT---HM 186
 Db 239 ALDIARMSYLLHN-RSKPITIHRDLKPRN--IIVDEHEHLKVGDFGLSKLIDVKLMDHYKM 296
 QY 187 TNNGSAAAMMAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPPDEIGGAPARIMAVNHG 246
 Db 297 TGGTGSYRYMAPEVEHQPDYKSDVYFSFGMIIYEMEGVAPFED--KQAYDAATLVARD 354
 QY 247 TRPPLIK--NLKPRIESLMTKRCWSKDPQSPRSMEEIVK 282
 Db 355 DKREPMRAQTYPPOMKALIEDCMSPTTKKPPFVEIVK 392

RESULT 13
 US-09-977-269-10
 Sequence 10, Application US/09977269
 Patent No. US20020082037A1

GENERAL INFORMATION:
 APPLICANT: ULRICH, AXEL
 APPLICANT: GISHIZKY, MIKHAIL
 APPLICANT: SURES, IRMINGARD
 TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
 FILE REFERENCE: 038602/1260
 CURRENT APPLICATION NUMBER: US/09/977,269
 CURRENT FILING DATE: 2001-10-16
 PRIOR APPLICATION NUMBER: 08/232,545
 PRIOR FILING DATE: 1994-04-22
 NUMBER OF SEQ ID NOS: 24
 SOFTWARE: Patentin Ver. 2.1
 SEQ ID NO 10
 LENGTH: 527
 TYPE: PRT
 ORGANISM: Mus sp.
 US-09-977-269-10

Query Match 12.1%; Score 363.5; DB 10; Length 527;
 Best Local Similarity 34.3%; Pred. No. 3.6e-15;
 Matches 92; Conservative 48; Mismatches 111; Indels 17; Gaps 9;

QY 30 EIDYKEIEVEEVGRGAFGVCKAKKRAK-DVAIKOI-ESESEKAFIVELRQLSRVNHP 87
 Db 260 EINSSELTFRBELSGSLFGVVRIGKRAQYKAIKAIREGAMCEEDFIEBAKVMKLTHP 319
 QY 88 NIVKLYGACL--NPVCLVMEYAEAGSLYNYLHGAEPPLPYTTAAHMSWCLQSGQVAVYH 145
 Db 320 KLVOLYGVCTQOKFIYVTFEMERGCLNPLRQ--GHSRMLLSWCDVCEGMEYL- 376
 QY 146 SMOKALIHRLDKPNNLLVAGCTVLYKICDQGT---CDIQTMTNNGSAAAMMAPEV 201
 Db 377 --ERNSTFHRDLARNCVLVEAG--VAVYSDFGMARVYLDQYTSSSGAKFPVKMCEPEV 433
 QY 202 EGSYSEKCDVFSMGIIIMEVITR-RKPFDEIGGAPARIMAVHNQTRPPLIKNLPRP 260
 Db 434 NYSRFSKSDVYSGVILMELIFTEGRMPFEK--NTNVEVVTWYTRGHLRHPKLVATKYLY 491
 QY 261 SLMTKRCWSKDPQSPRSMEEIVKIMTHLM 288
 Db 492 EVMDCWQERREGRSPFEDLRTIDELV 519

RESULT 14
 US-09-882-166-4
 Sequence 4, Application US/09882166
 Patent No. US20020151005A1

GENERAL INFORMATION:
 APPLICANT: Meyers, Rachel A.
 TITLE OF INVENTION: 53070, A NOVEL HUMAN PROTEIN KINASE
 TITLE OF INVENTION: FAMILY MEMBER AND USES THEREOF
 FILE REFERENCE: 10448-067001
 CURRENT APPLICATION NUMBER: US/09/882,166
 CURRENT FILING DATE: 2001-06-15
 PRIOR APPLICATION NUMBER: 60/212,078
 PRIOR FILING DATE: 2000-06-15
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 4
 LENGTH: 277
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: consensus sequence
 US-09-882-166-4

Query Match 12.0%; Score 360.5; DB 10; Length 277;
 Best Local Similarity 35.6%; Pred. No. 2.6e-15;

Job time : 119.014 secs

Matches 105; Conservative 42; Mismatches 81; Indels 67; Gaps 13;

Qy	37	EVEEIVGRGAFGVVCAKWR - AKDVAIKQIESSESRKAFIVELRQLSRVNHNPVVKLY 96
Db	2	ELJUEKGEFGVKYKAKHTGKIVAKILUKESLR ---EQILRKRUSHPNIVRELLGV 57
Qy	96	CLNP ---VCILVMEYBEGGSLYNLVHGAEPYPPYTAAHAMSWCICQSQGVAYLHSMQFAL 152
Db	58	FEDTDHLYLVMEYMEGGDLFDVLRNRGPL ---SEKEAKKIALQILRGLEYLHS ---NGI 111
Qy	153	IHRDLKPPNLLLVAGGTIVLKI CDFTGACDITQTHMTNNKGSAAW -MAPEV -PFGSNYSSEK 209
Db	112	VHRDLKPENLLDENGTV -KIADFGLA -RLLEKLTTPVGT -PWYMAPEVILEGREGYSSEK 168
Qy	210	CDVFSGIIILWEVI -----TRRKPFE 231
Db	169	VDVMSLVGVIYELLTGGLPFGADLPFAFTGQDEVQDLIFVLKLPFSDDELPKTRIDPLEE 228
Qy	232	IGGPAFRIMVAVNGTRPPIIKNLPKPIESLMTRCWSKDPSPQSPMBEEIVKIMTH 286
Db	229	L ----PRI ----KKRRLPLPNSCSBELKDLKKLNDKPSKPGGATAKEILNH 274

```

RESULT 15
US-09-842-582-4
; Sequence 4, Application US/09842582
; Patent No. US20020155570A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 2246, NOVEL PROTEIN KINASE MOLECULES AND
; TITLE OF INVENTION: USES THEREFOR
; FILE REFERENCE: 38155-20054.00
; CURRENT APPLICATION NUMBER: US/09/842,582
; CURRENT FILING DATE: 2001-04-25
; PRIOR APPLICATION NUMBER: US 60/199,391
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-09-842-582-4

```

Qy	37	EVEWGRGAFGVVCKAKWR--AKDVALKQIETESSERKAFIVELQRLSRVNHPIVKLYGA	95
Db	2	ELLEKLGEFGKVKYKAHKTGTVAVKILKES--LSLREIQILKRLSHPIVRLLGV	58
Qy	96	CLNP---VCLVMEYAEGGSYLVNLHGAEPIFYPTAAHAMSWCQCOSGVAYLHSMQPKAL	152
Db	59	FEDDDHLYLVMEYMGDDLFDYLRNGPL---SEKAKKIAQLRLGLEYLHS---NGI	112
Qy	153	IHRDKPNLLVAGGTVVKLICDGTGACDIQTHMNNKGSAAW--MAPEV-FGSGSYSEK	209
Db	113	VHRDLKPNILLDENGTV-KIADFGLA-RLLEKLTTFVGT-PWYMAPEVILEGSGYSSK	169
Qy	210	CDVFSWGIIILWEVI-----TRRKPFE	231
Db	170	VDVWSLVIIYELLTGGLPPGADLPFTGGDEVQDQLIFVLKLPFSDLPKTRIDPLEE	229
Qy	232	IGGPAFRIMAVHNGTRPPIKLNLPKPIESLMTKCSWKSPQRSMBEEIVKIMTH	286
Db	230	L-----KRRLLPSPNCSBELLKLLKLNKDPKSPGSAKEILNH	275

Search completed: December 9, 2002, 23:04:25

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: December 9, 2002, 22:58:34 ; Search time 3633.32 Seconds
(without alignments)
4637.778 Million cell updates/sec
Title: US-09-830-144-2
Perfect score: 3014
Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKQLEVRSSQKKRQGTS 579

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlp
-Q/cgn2.1/USPTO spool/US09830144/runat_04122002_141353_2252/app_query.fasta_1.1422
-DB=GenEmbl -OPMT=fastap -SURFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @cgn_1.1 5173 @runat_04122002_141353_2252 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.man.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3006	99.7	2757	9	BC017715 Homo sapi
2	3006	99.7	2769	6	AX377912 Sequence
3	3006	99.7	2769	9	AB009356 Homo sapi
4	3006	99.7	2785	6	E38397 NF-kappa B
5	2986	99.1	2443	10	MUSTAK1 D76446 Mouse mRNA
6	2982.5	99.0	2850	9	AB009357 Homo sapi
7	2982.5	99.0	2866	6	E38398 NF-kappa B
8	2963.5	98.3	3107	10	BC006665 Mus muscu
9	2756.5	91.5	1745	9	AF218074 Homo sapi
10	2746	91.1	1704	6	E38399 NF-kappa B
11	2729	90.5	1705	9	AB009358 Homo sapi
12	2692	89.3	135147	2	AC114407 Mus muscu
13	2638.5	87.5	2812	5	XLU92030 Xenopus lae
14	1476	49.0	3482	9	HSM800550
15	971	32.2	3349	3	AF159466 Homo sapi
16	971	32.2	3386	3	AY051933 Drosophil
17	910	30.2	2213	3	AK055901 Homo sapi
18	702.5	23.3	32564	2	AC014558 Drosophil
19	702.5	23.3	164942	3	AC011758 Drosophil
c 20	702.5	23.3	302303	3	AE003571 Drosophil
c 21	502	16.7	3138	9	HSMSTM
22	493.5	16.4	3454	6	AX337846 Sequence
23	493.5	16.4	3454	9	HARNAMUK2 X90846 H.sapiens m
24	490	16.3	2120	6	AR119790 Sequence
25	490	16.3	2120	6	AR126750 Sequence
26	490	16.3	2120	6	AR128910 Sequence
27	490	16.3	2120	6	AR130841 Sequence
28	490	16.3	2120	6	AR138886 Sequence
29	490	16.3	2120	6	AR141354 Sequence
30	482	16.0	1370	9	AF325454 Homo sapi
31	482	16.0	1571	9	BC001401 Homo sapi
32	482	16.0	2251	9	AK056310 Homo sapi
33	482	16.0	7195	9	AF480462 Homo sapi
34	481	16.0	1365	6	AR119791 Sequence
35	481	16.0	1365	6	AR126751 Sequence
36	481	16.0	1365	6	AR128911 Sequence
37	481	16.0	1365	6	AR130842 Sequence
38	481	16.0	1365	6	AR138887 Sequence
39	481	16.0	1365	6	AR141355 Sequence
40	481	16.0	1368	9	AB049734 Homo sapi
41	481	16.0	1429	10	AB049732 Mus muscu
42	480	15.9	3931	9	AF251442 Homo sapi
43	474	15.7	8125	3	AF481923 Dictyoste
44	462	15.3	3531	9	HSU07747 Human SH3 d
45	462	15.3	3558	6	AX399680 Sequence

ALIGNMENTS

RESULT 1

BC017715	LOCUS	2757 bp	mRNA	linear	PRI 06-DEC-2001
DEFINITION	Homo sapiens, mitogen-activated protein kinase kinase 7, clone MGC:21263 IMAGE:3906837, mRNA, complete cds.				
ACCESSION	BC017715				
VERSION	BC017715.1				
KEYWORDS	GI:17389342				
SOURCE	MGC.				
ORGANISM	Homo sapiens.				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	1 (bases 1 to 2757)				
TITLE	Strausberg,R.				
JOURNAL	Direct Submission				
REMARK	Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
COMMENT	NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgaps-rc@mail.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www.sbgc.stanford.edu Contact: (Dickson, Mark) mcdpaxil@stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.				
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAX Plate: 22 Row: 1 Column: 18 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507360.				
source	Location/Qualifiers				
	1..2757				
	/organism="Homo sapiens"				
	/db_xref="LocusID:6885"				
	/db_xref="taxon:9606"				
	/clone="MGC:21263 IMAGE:3906837"				
	/tissue.type="uterus", "leiomyosarcoma"				
	/clone_id="NH_MGC_71"				
	/lab_host="DH10B"				
	/note="Vector: pCMV-SPORT6"				
	198..1937				
	/codon_start=1				
	/product="mitogen-activated protein kinase kinase 7"				
	/protein_id="AAH17715.1"				
	/db_xref="GI:17389343"				
	/translation="MSTASAASSSSSSAGMTAPSPVNFPEIDYKKEIFVEVGRGAFVACAKWRADVAIKQIESERKAFIVELRQSRVNHNTVTKYGCALNPVCLVMEYAEGLSLVNLHGAEPLDPIYTAHAMQSLQSGVAVLHMQPALIHRDLKPPNLLVAGGTVLKI CDGFACDIQHTMTNKGSAAMAPVEVGGSNYSBKCDVFSWGII LMEVITRRKRPDEIGGPAFRIMVAOYCHVSDGQSNASTSGSPMDIASNTNSKSDINMEIYKINTHLMRYPPGADPELTPCYPCOVSDDGQSNASTSGSPMDIASNTNSKSDINMEOVPAINDTIRLESKILKNDAKQSSSGRLSAGSGSVEVSLPTSGEGRMSADMSEIRAIATTGNGGPRRRSIODLTVTGTBEQGVSSRSSTSVKMITTSGTTSKAPTRSSHPMTDDSTGNGSDNSIIPMAVLTLDQLPLAVCPNKSRSMAVFEQHCMAQEVNIMKVTETALLRKQELVAELDDQEKQOQNTSLVDEHKLLDENKLSLSTYYQCKRQKLEL VTRSQDQKRQKTS"				
BASE COUNT	819 a 566 c 647 g 725 t				
ORIGIN					
Alignment Scores:					
Pred. No.:	4.88e-177 Length: 2757				
Score:	3006.00 Matches: 578				
Percent Similarity:	99.83% Conservative: 0				
Best Local Similarity:	99.83% Mismatches: 1				
Query Match:	99.73% Indels: 0				
DB:	9 Gaps: 0				

[illegible]

Db	343	GCTATTAAACCAATAGAAAGTGAATCTGAGAGGAAAGCGTTTATTGTAGAGCTTCGGCAG	402
Qy	81	LeuSerTrgValAsnHisIspRobsnIleValIysLeuTrgValaCysLeuAsnProVal	100
Db	403	TTTATCCCGTGTGAACCATCTCTAATATTGTAAGCGTTTATGAGAGCTTCGTAATCCAGTG	462
Qy	101	CysLeuValMetGluTrgValaGluTrgValaGlySerLeuTrgAsnValLeuHisGlyValaGlu	120
Db	463	TGCTCTGTGTGAAATAGTCTGAAGGGGCGCTTTATATATATGCTTGCAAGTGGTCTGAA	522
Qy	121	ProLeuProTrgTrgThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	523	CCATTGCGCAATTTATCTGCTGCTGCCACGCAAGATTTGCTTTACAGTGTCCCAAGG	582
Qy	141	ValAlaTrgLeuHisSerMetGlnProValaLeuIleHisArgAspLeuLysProPro	160
Db	583	GTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTCCAGAGGACCTGAAACCAACA	642
Qy	161	AsnLeuLeuValAlaGlyGlyThrValaLeuValIleCysAspPheGlyThrAlaCys	180
Db	643	AACTACTGCTGGTGTGAGGGGGGAGAGTCTTAAATAATTTGTGATTTGGTACAGCTGT	702
Qy	181	AspIleGlnThrIsmetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	703	GACATTCAGACACATGACCAATTAACAAGGAGAGTGTCTGTGATGCGACCTGAAAT	762
Qy	201	PheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp	220
Db	763	TTTGAGAGTATGATTTACAGTGAATAATGACGCTTCAGCTGGGGATATATTCCTTTGG	822
Qy	221	GluValIleThrArgGlyLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	823	GAAAGATGAACGCTCGGAAACCTTTGATGAGATTGGTGGCCAGCTTTCGATATATG	882
Qy	241	TrpAlaValHisAsnGlyTrpArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	883	TGGGCTGTTCAATATGATCTGACACACACATGATAAAAATTTACTAACCCATTGAG	942
Qy	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	943	AGCCGATGACTCGTGTGGTCTAAAGATCCTTCCACGCGCTTCATGAGAGAAATT	1002
Qy	281	ValLysIleMetThrHisLeuMetArgTrpPheProGlyValaAspGluProLeuGlnTr	300
Db	1003	GTGAAATAAAGACTCACTGATGCGGTACTTTCCAGAGCAGATGAGCATTTACAGTAT	1062
Qy	301	ProCysGlnTrpSerAspGluGlyGlnSerAsnSerAlaThrSerTrgIleSerPheMet	320
Db	1063	CCCTTCAGTATTCATGATGAGAGACAGACATCTGCCACACAGTACAGGCTCATTTATG	1122
Qy	321	AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla	340
Db	1123	GACATTTGCTTCACAAATATACAGTAAACAAAGTGAACATTAATATGAGAGCAAGTCTG	1182
Qy	341	ThrAsnAspThrIleLysArgLeuGlnSerLysLeuLeuLysAsnGlnAlaLysGlnGln	360
Db	1183	ACAAATGATACTATTAAACGCTTGATCAAAATTTGTGAAAAATAGCAAAACCAAG	1242
Qy	361	SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro	380
Db	1243	AGTGAATTCGGACGTTTAAGCTTTGAGAGCTCCCGGTGGAGCAGTGTGAGAGCTTGCC	1302
Qy	381	ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla	400
Db	1303	CCAACCTCGAGAGGCAAGAGTATGATGTCAATTAAGAACTTGAGTGCCTGACC	1362
Qy	401	AlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly	420
Db	1363	GCAACACAGCGCAAGGACAGCAAGACGTATGATCCATCCAAACTTGACTTAATCTGA	1422
Qy	421	ThrGluProGlyGlnValSerSerArgSerSerSerProSerValaArgMetIleThrThr	440
Db	1423	ACGAACCTCGTACAGTAGACGTAGGTCACTCCAGTCCAGATGATGATTTACAC	1482

Oy	441	Seq1yProthriserGluysProthraSgerHisProTyrThProaSpaSerThr	460
Db	1483	TCAGGACCAACCTCGAAAAGCCAACTCCGAAGTCATCCATGAGCCCTGATGATTCACA	1542
Oy	461	AspThraSnglyserApaSasSer11eProMetAlTyLeuThrLeuAspHsgInleu	480
Db	1543	GATACCAATGATACAGTATACCTCATCCCATGGCTTACTTACATGGATCAACCACTA	1602
Oy	481	GlnProleuAlaProCysProaSenSerlyGluSerMetAlaValPheGluGlnHisCys	500
Db	1603	CAGCTCTAGACACCGTCCCAAACTCCAAAGATCTTATGGCAGTGTTTGAACAGCATTTGT	1662
Oy	501	LysMetAlaGlnGluTyrmetyLysValGlnThGlu11eAlaLeuLeuGlnGlnArglys	520
Db	1663	AAATGGCCACAGAAATATATGAAGTTCAACAGAAATTCATGTCATGTTATACAGAAAG	1722
Oy	521	GlnGlnLeuValAlaGlnLeuAspGlnAspGlnLysAspGlnGlnGlnHisThrSerArgLeu	540
Db	1723	CAAGACTAGTTGCAGAACTGCACAGAGATGAAAGACGACGAAATATCATCTCCCTCG	1782
Oy	541	ValGlnGlnHisLysLysLysLeuLeuAspGlnLysSerLeuSerThrTyrTyrglnGln	560
Db	1783	GTACAGGAACCTAAAGAGCTTTTGAATGAAGAAACAAAGCCCTTCTACTTACTTCACGCAA	1842
Oy	561	CysLysLysGlnLeuGlnVal11eArgSerGlnGlnGlnLysArgGlnGlyThrSer	579
Db	1843	TGCAGAAAACAACTAGAGCTCATCAGAACTCAGAGCAGAAAGACAGACCACTTCA	1899
RESULT 5			
MUSTAKI			
LOCUS		2443 bp	RNA
DEFINITION		Mouse mRNA for TAK1 (TGF-beta-activated kinase), complete cds.	
ACCESSION	D76446		
VERSION	D76446.1	GI:1167505	
KEYWORDS	TAK1; TGF-beta-activated kinase; protein kinase.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Yanaguchi, K., Shirakabe, K., Shibuya, H., Irie, K., Oishi, I., Ueno, N., Taniguchi, T., Nishida, E. and Matsunoto, K.		
TITLE	Identification of a member of the MAPKK family as a potential mediator of TGF-beta signal transduction		
JOURNAL	Science 270 (5244), 2008-2011 (1995)		
MEDLINE	96123277		
REFERENCE	2 (bases 1 to 2443)		
AUTHORS	Matsunoto, K.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-OCT-1995) Kunihito Matsunoto, Faculty of Science, Nagoya University, Department of Molecular Biology; Furoo-chou, Chikusa-ku, Nagoya, Aichi 464-01, Japan		
FEATURES	(E-mail: g94117@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-3000, Fax: 052-789-3001)		
source	Location/Qualifiers		
	1..2443.		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	157..1896		
	/codon_start=1		
	/product="TAK1 (TGF-beta-activated kinase)"		
	/protein_id="BA011184.1"		
	/db_xref="GI:1167506"		
	/translation="MSTRSASSSSSSSSASSEMTEAPSOVUNFPEIDYKEIEVEIVCR		
	GAFVGECAKARADVAIKQIESSEKATFVEIRQSRVNHPIYALYACINPCL		
	VMVEGEGSLNVLDHGAEPYVYTAHANSMLQSGVAVLHSMOKALIHDLKP		
	NLIVAGGTVLKICDFGACADIQTHMNNKGSAAWMAVEFEGSNVSEKCDVSWGII		
	LMETVTRKKPDEIGGPAFRIMAHVHGNTPEPLKNLPKPIESIMTCWGSQSRSS		
	MEIYKINTHLMRYPRPADDELPYPCOYSGRGSNSATISGINDISTNKSNDN		
	MEQVAINDITIKRLSKLKKQAKQOESGSLSGARGSSVSLPTSGSKNSDM		
	SEIEMVAVTAGNQPRRSIQDILVTGTEGVSSSSSVSMITTSGETSEKPR		
	SHPWTPDSDTJNDSNDSIPMAVYITLDHOLPLAPCPNSESMAVFQHKMAQEIYMK		

RESULT 6	
AB009357	
LOCUS	2850 bp mRNA linear PRI 04-MAR-1998
DEFINITION	Homo sapiens mRNA for TGF-beta activated kinase lb, complete cds.
ACCESSION	AB009357
VERSION	AB009357.1 GI:2924625
KEYWORDS	TAK1b; TGF-beta activated kinase lb.
SOURCE	Homo sapiens lung cDNA to mRNA, clone_lib:lambda gt11 clone:pBS7AK1b.
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (sites)
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE	TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an

Qy 61 AlaIleIysGlnIleGluSerGluSerGluArgIysAlaPheIleValGluIleuArgGln 80
 Db 328 GCTATTAAACAGATAGAAAGTGTCTGAGAGGAAGGCTTTCAATTGTGAGGCTCGGCGAG 387
 Qy 81 LeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCysLeuAsnProVal 100
 Db 388 TTGTGGGTGTGAACCATCTCACTTGTCAAGTTGTACGGAGCCTGCCTGAATCCAGTA 447
 Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValIleuHisGlyAlaGlu 120
 Db 448 TGTCTGTGTGATGAATATGACAGAGGGGGCTCAATTGTATTAATGTGTGATGTTGTGAA 507
 Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
 Db 508 CCATTGCTTACTACTGCTCATGCTCATGATGAGCTGTGTGTACAGTTGTCCCAAGGA 567
 Qy 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
 Db 568 GTGGCTTACCTGCACAGCATGACGCCAAAGCGCTGATTACAGGGACCTCAAGCCCTCCA 627
 Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 Db 628 AACTGTGCTGGTGTGAGGAGGACAGTTCTAAAAATCTGCCATTGTGTACAGCTTGT 687
 Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
 Db 688 GACATCCAAAACACACATGACCAATAATAAAGGAGTCTGCTTGGATGGCCCTGAAGTG 747
 Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTyr 220
 Db 748 TTTGAAGGTAGCAATTAACGTGAAGTGTGATGCTTTCAGCTGGGGGTATTTATCCTCTGG 807
 Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
 Db 808 GAAGTGAATAACCGCGGAAACCCCTTCATGAGATCCGTGGCCAGCTTTCAGAAATCATG 867
 Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
 Db 868 TGGGCTGTTCATATGTCACCTCGACCCACTGATCAATAAATTTACCTAAGCCCATTTAG 927
 Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
 Db 928 AGCTTGATGACAGCTGTGTGTCTAAGACCCATCTCAGCGCCCTTCAATGGAGGAAT 987
 Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
 Db 988 GTGAATAATAAGTCACTTCTGATGGGTACTTCCAGGAGCGGATGAGCCATTACAGTAT 1047
 Qy 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
 Db 1048 CTTTGTCTAGTACTCTGATGAAGGGCAGAGCAACTCAGCCACCCAGCACAGGCTCGTTTCATG 1107
 Qy 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
 Db 1108 GACATTGCTTCTACAAATACCAAGTAAATAAGTGACACAATAATATGGACAGGTTCTGTGCC 1167
 Qy 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
 Db 1168 ACAACGACACTATTAAACGCTTGAGTCAAAACTGTGTAAAAAACCCAGGCAAGCAACAG 1227
 Qy 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
 Db 1228 AGTGAATCTGGACGCTGAGCTTGGAGGCTCTCTGTGGAGCAGTGTGGAGAGCTTGCCC 1287
 Qy 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
 Db 1288 CCCACTTCGAGGGCAAGAGGATGATGCTGACATGCTGAAATAGAAAGCCAGGATCGTG 1347
 Qy 401 AlaThr----- 402
 Db 1348 GCGACTGCGACGCTTATCCAAAGCCTAAACGGGGGCCACCGTAAAAACCGCTTCAATTGGCAAC 1407
 Qy 403 -----ThrGlyAsnGlyGlnProArgArgArgSerIle 413

Db 1408 ATTCTGATGTCCTCGATCGTCATATCAGGTAAACGGCAACCAAGCGGTAGATCCATC 1467
 Qy 414 GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerPro 433
 Db 1468 CAAGACTTGACTGTACTTACTGGGACAGAACTCTGTTCAGGTGAGCAGCGGTATCCAGCCCT 1527
 Qy 434 SerValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisPro 453
 Db 1528 AGTGTCAAGATGATCACTACCTCAGGACCAACCTCAGAAAGCCAGCTCGCAGTCACCCA 1587
 Qy 454 TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr 473
 Db 1588 TGGACCCCTGATGATTCACAGACACCAATGGCTCAGATAACTCCATCCCAATGGCGTAT 1647
 Qy 474 LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet 493
 Db 1648 CTTACACTGGATCACCAGCTACAGCCTCTAGCGCGCTGCCAAACTCCAAAGATCCCATG 1707
 Qy 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
 Db 1708 GCAGTGTTCGAAACAGCACTGTAAAAATGGCACAGGAGTATATGAAGTTCAAACCGAAATC 1767
 Qy 514 AlaLeuLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGlnAspGluLysAsp 533
 Db 1768 GCATTGTACTACAGAGAAAGCAAGAACTAGTTGCCAGAAATTGGACCAGGATGAAAGGAC 1827
 Qy 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSer 553
 Db 1828 CAGCAAAATACATCTCTGCTGTGTACAGGAACATAAAAAAGCTTTTAGTAAAAACAAAGC 1887
 Qy 554 LeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnGln 573
 Db 1888 CTTTCTACTTATTACAGCAATGCAAAAAACAACTAGAGGTTCATCAGAAAGCCCAACAGCAG 1947
 Qy 574 LysArgGlnGlyThrSer 579
 Db 1948 AAACGACAGGCACTTCA 1965
 RESULT 9
 AF218074 1745 bp mRNA linear PRI 18-DEC-2000
 LOCUS Homo sapiens TGF beta-activated kinase splice variant d (TAK1)
 DEFINITION
 mRNA, complete cds.
 ACCESSION AF218074
 VERSION AF218074.1 GI:6746614
 KEYWORDS
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1745)
 AUTHORS Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.
 TITLE Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1
 JOURNAL Biochim. Biophys. Acta 1517 (1), 46-52 (2000)
 MEDLINE 20568765
 PUBMED 11118615
 REFERENCE 2 (bases 1 to 1745)
 AUTHORS Dempsey,C.E. and Guesdon,F.
 TITLE Direct Submission
 JOURNAL Submitted (21-DEC-1999) Division of Molecular and Genetic Medicine,
 University of Sheffield, School of Medicine, Glossop Road,
 Sheffield S10 2UP, United Kingdom
 FEATURES
 source 1. .1745
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 gene 1. .1745
 /gene="TAK1"
 CDS 7. .1482
 /gene="TAK1"
 /note="TAK1d"

```

/codon_start=1
/product="TGF beta-activated kinase splice variant d"
/protein_id="AAP27652.1"
/db_xref="gi:6746615"
/translation="MSTASASSSSSSAGBMTIAPSOVLNPFEDIDKEIEVEVNGR
GAEVVCAGKAKRAKVAIKQIESSEKAFIVELROLSRNHNHYKLYACINPCL
VMEYAEGLSYNLVHGAEPLEPYTAHAMSVCLOCSGVAYLHSMOKALINHDLP
NLIVAGTVDLKI CD FGTACDI QHMTNNKSAAMAPEVEGSENSEKDVSWGI
LMEVITRRKPFDEIGPAPRIMAVHNGTRPPLIKNLPKPIESIMTRCWSKDSQBS
MEJIKMTHTMRYEPGADPELOYPCOYDEGNSATSTGSEMDIASTVSNKSDTN
MEQVPATNDTIRLESKLIKNOAKOSESGLSIGASRGSVESLPTSGKMSADM
SEIARIATTTGNGQPRRSIQDITVGTGTEPGQVSSSSPSVRMITTSPTSEKPTR
SHPWTPDSDTNGSDNSIPMAVLTLDHQIARTSCRTGPE"

```

BASE COUNT 526 a 375 c 412 g 432 t
ORIGIN

Alignment Scores:

	Pred. No.:	7.79e-162	Length:	1745
Score:	2756.50	Matches:	538	
Percent Similarity:	92.93%	Conservative:	1	
Best Local Similarity:	92.76%	Mismatches:	1	
Query March:	91.46%	Indels:	40	
DB:	9	Gaps:	1	

US-09-830-144-2 (1-579) x AP218074 (1-1745)

```

QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGluMetIleGlu 20
DB 7 ATGTCACAGCCTTCGCCCTCCTCCTCCTCCTGCGCGCGAGATGATCGAA 66
QY 21 AlaProSerGlnValLeuAsnPhelGluIleAspTyrLysGluIleGluValGluGlu 40
DB 67 GCCCTCCAGGTCCTCAACTTTGAAAGATGACTACAAAGAGATCAGAGTGAAGAG 126
QY 41 ValValGluArgGluValAlaPhelGluValAlaGluValAlaGluValAlaGluVal 60
DB 127 GTTGTGGAGAGGAGCCCTTGGAGTTGTTGCAAAAGCTAAGTGAAGCAAAAGATT 186
QY 61 AlaIleGluGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 187 GCTATTAAACAATAAAGATGATGAGAGAAAGCGTTATTATGAGCTTCGCGAG 246
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGluValAlaGluLeuAsnProVal 100
DB 247 TTATCCCGGTGAACCATCTTAATTGTAAGCTTTATGAGCCGTGTAATCCAGTG 306
QY 101 CysLeuValMetGluTyrValAlaGluGluGluSerLeuTyrAsnValLeuHisGluValGlu 120
DB 307 TGCTTGTGATGGAATAGCTGAAAGGGGCTCTTTATATATGTGCTGATGCTGAA 366
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 367 CCATTGCCATATATATCTGCTGCCAGCAATAGTTGGGTGTTACAGTGTCCCAAGA 426
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 427 GTGGCTTATCTTCACAGCATGCAACCAAGCCCTTAATTCACAGGAGCTTGAAACCA 486
QY 161 AsnLeuLeuLeuValAlaGluGluGluTyrValLeuLysIleCysAspPheGluThrAlaCys 180
DB 487 AACTTACTGCTGTGAGGGGGAACGTTCTAAAAATTTGTGATTTTGTACAGCTCT 546
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGluSerAlaAlaTyrMetAlaProGluVal 200
DB 547 GACATTTCAGACACACATGACCAATTAACAAGGAGTGTCTTGATGGCACTGAAGTT 606
QY 201 PheGluLysSerAsnTyrSerGluLysCysAspValPheSerTyrGluIleIleLeuTyr 220
DB 607 TTGGAAGGTAGTAAATACAGTGAATAATGAGCTTTCAGCTGGGATATATCTTTTGG 666
QY 221 GluValIleThrArgArgLysProPheAspGluIleGluGluProAlaPheArgIleMet 240
DB 667 GAAGTGAATACGCGTCGAAACCTTTGATGAGATTGTGGCCACGCTTCCAAATCAAG 726

```

```

QY 241 TTPAlValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 727 TGGCGCTTCATATATGATCTCGACCAACGATPAAAAATTTACTTAAAGCCATTGAG 786
QY 261 SerLeuMetThrArgCysTyrPheSerLysAspProSerGlnArgProSerMetGluGluIle 280
DB 787 AGCTTATGACTGTTGTTGTGTAAAGATCTCTCCAGCGCCCTTCAATGAGGAAAT 846
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGluValAspGluProLeuGlnTyr 300
DB 847 GTAAAAATATGACTCATTGATGCGGTACTTCCAGAGCAGATGAGCCATTACAGTAT 906
QY 301 ProCysGlnTyrSerAspGluGluGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
DB 907 CCTGTGATGATTAATGATGAAGACAGACCACTGCGCAACGATACAGCTCATTCATG 966
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
DB 967 GACATTGCTCTTCAAAATACAGATTAACAAAGTACACTAATATGAGCAAGTCTGCGC 1026
QY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
DB 1027 ACAATGATACTATTAAAGCGCTTAAATCAAAATTTGTGAAAAATCAGGCAAGCAACAG 1086
QY 361 SerGluSerGluArgLeuSerLeuGluValAlaSerHisGlySerSerValGluSerLeuPro 380
DB 1087 AGTGAATCTGACGTTAAAGCTTGAAGCTCCCGTGGAGCGAGTGTGAGAGCTTGCCC 1146
QY 381 ProThrSerGluGluLysArgMetSerAlaAspMetSerGluIleGluValArgIleAla 400
DB 1147 CCAACCTCTAGAGGCAAGAGATGATGCTGACATCTTGAAATTAACCTAGATGAGTGC 1206
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
DB 1207 GCAACCAACAGCAACGACACCAAGACGTATCCATCCAACTGACTGACTGATCTGAA 1266
QY 421 ThrGluProGluGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
DB 1267 ACAGAACCTGTGACGGACAGATGATGATCCAGTCCAGTCCAGATTCAGATTAATCAC 1326
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
DB 1327 TCAGGACCAACCTAGAAAGCCAACTCGAATGCATCCATGAGCCCTGATGATTCACA 1386
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
DB 1387 GATACCAATGATCAGATTAATCTCCATCCCAATGGCTTATCTTACCTGATCCCAACTA 1446
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
DB 1446 ----- 1446
QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuGlnArgLys 520
DB 1447 -----CAG 1449
QY 521 -GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLe 540
DB 1450 GCAAGAACTAGTTGCAGAACTGAGCCAGATGAAAGACACAGCAAAATCACTCGCCT 1509
QY 540 ValAlaGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnG 560
DB 1510 GGTACAGGAACAATAAAAGCTTTAGATGAAACAAAAGCCCTTTCTTACTTACACAGCA 1569
QY 560 nCysLysLysGlnLeuGluValAlaArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
DB 1570 ATGCAAAATAACACTAGAGTCAACAGATCAGACACAGAAACGACAGGCACTTCA 1627

```

RESULT 10
E38399 1704 bp DNA linear PAT 31-JAN-2002
LOCUS NF-kappa B activation inhibitory drug targeting TAK1 and method for
DEFINITION

QY	474	LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerGlySerMet	493
Db	1501	CTTACACTGGATCACCACTACAG-----	1524
QY	494	AlaValPheGlnGlnHisCysLysMetAlaGlnGlnIuTyMetLysValGlnThrGluIle	513
Db	1524	-----	1524
QY	514	AlaLeuLeuLeuGlnAlaGlyGlnGlnIuLeuValAlaGlnIuLeuAspGlnAspGlnLysAsp	533
Db	1525	-----CAAGAACTAGTTGGAGAACTGGACACAGATGAAAGGAC	1563
QY	534	GlnGlnAsnThrSerArgLeuValGlnGlnIuHisLysLysLeuLeuAspGlnAsnLysSer	553
Db	1564	CAGCAAAATACATCTTCGCTGGTACAGAACTATAAAAGCTTTTGATGAAACAAAGGC	1623
QY	554	LeuSerThrTyTyTyGlnGlnCysLysLysGlnLeuGlnValIleArgSerGlnGln	573
Db	1624	CTTTCTACTTACTTACAGCAATGCAAAAAACAATAGAGTCATCAGAAGTCAGCAGCAG	1683
QY	574	LysAlaGlnGlnIuThrSer	579
Db	1684	AAACGACAAAGCCTTCA	1701
RESULT 11			
LOCUS	AB009358	1705 bp	linear
DEFINITION	Homo sapiens mRNA for TGF-beta activated kinase 1c, complete cds.		PRI 20-JAN-2001
ACCESSION	AB009358		
VERSION	AB009358.2	GI:8978251	
KEYWORDS	TAK1c; TGF-beta activated kinase 1c.		
SOURCE	Homo sapiens cell_line:Hela cDNA to mRNA, clone:pT7TAK1c.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.		
TITLE	TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism		
JOURNAL	Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)		
MEDLINE	98153801		
REFERENCE	2 (sites)		
AUTHORS	Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.		
TITLE	Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1		
JOURNAL	Biochim. Biophys. Acta 1517 (1), 46-52 (2000)		
MEDLINE	20568765		
REFERENCE	3 (bases 1 to 1705)		
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K., Sugita,T. and Guesdon,F.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-DEC-1997) Francois Guesdon, University of Sheffield, Royal Hallamshire Hospital, Division of Molecular and Genomic Medicine, Functional Genomics Group, Glossop road, Sheffield S10 2JF, United Kingdom (E-mail:f.a.guesdon@sheffield.ac.uk, Fax:44-114-271-3846)		
COMMENT	On Jul 8, 2000 this sequence version replaced gi:2924627.		
FEATURES	Location/Qualifiers		
source	1..1705		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="pT7TAK1c"		
	/cell_line="Hela"		
	1..1557		
	/note="TAK1c"		
	/codon_start=1		
	/product="TGF-beta activated kinase 1c"		
	/protein_id="BAA25027.2"		
	/db_xref="GI:8978252"		
	/translation="MSTASAASSSSSSAGEMITAPSOVLNPEIDYKEIFEVEYVGHG		
	CARFVYCAKRAKVAIKQISESEKRAFIYELRQLSRVNHPNIVKLYGACLNQPC		
	VNEVAGSGSLYVHGAEPLPYTYAHAAHSWQCSGGAYVSHNSQPALIHRDLKPI		
	NLLVAGSGVLKICDGTACTAQTQTHMTNKGSAAMABVEGNSYSEKCVFSGGI		

BASE COUNT	511 a	381 c	402 g	411 t
ORIGIN	ARTSCTGPG ⁹			
Alignment Scores:				
Pred. No.:	3,81e-160			
Score:	2729.00			
Percent Similarity:	88.63%			
Best Local Similarity:	90.47%			
Query Match:	98.54%			
DB:	9			
	Gaps: 2			
US-09-830-144-2 (1-579) x AB009358 (1-1705)				
QY	1	MetserThraAlaSerAlaAlaSerSerSerSerSerSerAlaGluMetIleGlu	20	1705
DB	1	ATGCTCAAGCCTTGCGCGCTCCCTCCCTCGTTCGCGCGGAGATGATCA	60	537
QY	21	AlaProSerGluValIleAsnPheGluIleAspTyrLysGluIleGluValGlu	40	1
DB	61	GCCCCCTCCAGGCTCCTCACTTGAGAGATCCACAGAGATGAGTGAAG	120	2
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTPArgAlaLysAspVal	60	67
DB	121	GTTGTGGAAGAGGAGCCTTGAGAGTGTTCGAAAGCTAGTGAAGCAAAAGATGT	180	2
QY	61	AlaIleLysGluIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80	2
DB	181	GCTATTAAACAAATGAAGAAGTCTGAAGAGGAAGCGTTATTGAGACCTTGCGAG	240	2
QY	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100	2
DB	241	TTATCCCGTGTGAACCATCTTAATTGTTAAAGCTTTATGAGCCTGTTGAATCAGTG	300	2
QY	101	CysLeuValMetGluTyrAlaGluGlyLysSerLeuTyrAsnValLeuHisGlyAlaGlu	120	2
DB	301	TGCTGTGATGGAATATGCTGAAGGGGGCTCTTATATTAATGAGCTCATGTGCTGAA	360	2
QY	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTProCysLeuGlnCysSerGlnGly	140	2
DB	361	CCATTGCCATTTATACGCTGCCACCAATGAGTTGTTACAGTGTTCCTCAAGA	420	2
QY	141	ValAlaTyrLeuHisSerMetGluProLysAlaLeuIleHisArgAspLeuLysProPro	160	2
DB	421	GTTGGCTTATCTTCACAGCATGCAACCCAAAGCGTTATTCACAGGAGACTGAAACACCA	480	2
QY	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180	2
DB	481	AACTACTAGCTGGTGGAGGGGGGACAGTTCTAAAAATTTGGATTGGTACACCTGT	540	2
QY	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaIleTProMetAlaProGluVal	200	2
DB	541	GACATTTCAGACACATATACCAATTAACAAGGAGAGTGTGCTGTGATGGCACTAAAGTT	600	2
QY	201	PheGluLysSerAsnTyrSerGluLysCysAspValPheSerTArgIleIleLeuTyr	220	2
DB	601	TTTGAAGGAGTAAATTACAGTGAAGAAAGTGAAGTCTTCAGCTGGGAGTATTTCTTTGG	660	2
QY	221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240	2
DB	661	GAAAGCATTAACCGTCGGAACCTTTATATAGATTGGTGGCCGAGCTTCCGAATCATG	720	2
QY	241	TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260	2
DB	721	TGGGCTGTTCATTAATGAGTACGACACCACTGATTAATAAATTTTACCTTAAGCCATTGAG	780	2
QY	261	SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle	280	2

Db	781	AGCGTCGATGACTCGTGTGTGTTCTAAAGATCCTTCCAGCGCCCTTCAATGAGGAAATT	840
Qy	281	VallysileMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	841	GTGAATAATGACTCACTTGATCGGTACTTTCCAGAGCAGATGAGCCATTACAGTAT	900
Qy	301	ProCysGlnTyrSerAspGluGlnSerAsnSerAlaThrSerThrGlySerPheMet	320
Db	901	CCTTGTGTCAGTATTACAGATGAAGGACAGAGCAACTCTGCCACCGATACAGGCTCATTATG	960
Qy	321	AspileAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla	340
Db	961	GACATTGCTTCTCAAAATACAGATGAACAAAAGTCACACTAATATGGAGCAAGTTCTCTGCC	1020
Qy	341	ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln	360
Db	1021	ACAAATGATATTAATAGCGCTTAGAATCAAAATGTTGAAAATCAGGCAAGCAACAG	1080
Qy	361	SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro	380
Db	1081	AGTGAATCTGGACGCTTAAAGCTTGGAGCCTCCCGTGGAGCAGTGTGGAGAGCTTGCCC	1140
Qy	381	ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla	400
Db	1141	CCAACTCTGAGGCAAGAGATGAGTGTGACATGTCTGAAATAGAAAGCTAGGATCGCC	1200
Qy	401	AlaThrThr-----	403
Db	1201	GCAACCAGCGCTATTCCAAGCCTAAAGGGGGCCACGTTAAACTGCTTCATTGGCAAC	1260
Qy	404	-----GlyAsnGlyGlnProArgArgArgSerIle	413
Db	1261	ATTCTGGATGTCCTGAGATCGTCATATCAGGCAACGACAGCAACGCTAGATCCATC	1320
Qy	414	GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerPro	433
Db	1321	CAAGACTTGACTGTAACCTGGAAACAGAACCTGGTCAGGTGACGATAGGTGCATCGATCCC	1380
Qy	434	SerValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisPro	453
Db	1381	AGTGTCAGATGATTAATCTCTCAGGACCAACCTCAGAAAAGCCAACTCGAAGTCATCCA	1440
Qy	454	TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr	473
Db	1441	TGGACCCCTGATGATTCACAGATACCAATGGATCAGATCAACTCCATCCCAATGGCTTAT	1500
Qy	474	LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet	493
Db	1501	CTTACACTGGATCACCAACTA-----	1521
Qy	494	AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle	513
Db	1521	-----	1521
Qy	514	AlaLeuLeuLeuGlnArgLys-GlnGluLeuValAlaGluLeuAspGlnAspGluLysAs	533
Db	1522	-----CAGGCAAGAAGTCTAGTTGCAGAACTGGACACGATGAAAGGA	1563
Qy	533	pGlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSe	553
Db	1564	CCAGCAAAATACATCTCGCTGGTACAGGAACATAAAAGCTTTTAGATGAAAAACAAGG	1623
Qy	553	rLeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnG	573
Db	1624	CCYTTCTACTTACTACAGCAATGCAAAAAACAACTAGAGGTTCATCAGAAAGTCAGCAGCA	1683
Qy	573	nLysArgGlnGlyThrSer	579
Db	1684	GAACGCAAGGCACTTCA	1702

RESULT 12

AC114407

LOCUS

AC114407 135147 bp DNA linear HTG 26-JUN-2002

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Mus musculus clone RP23-51G1, WORKING DRAFT SEQUENCE, 5 ordered pieces
AC114407.3 GI:21592111
HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
house mouse.

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 135147)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,

Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,

Kamat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R.,

Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,

Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,

McEwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,

Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,

Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,

Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,

Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,

Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuback,R., Seaman,S.,

Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,

Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,

Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,

Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,

Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 135147)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,

Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,

Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,

Cook,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S.,

Faro,S., Ferreira,P., FitzGerald,M., FitzHugh,W., Gage,D.,

Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L.,

Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I.,

Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRocque,K.,

Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K.,

Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N.,

Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrim,J.,

Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,

Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P.,

O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,

Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S.,

Schuback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N.,

Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S.,

Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J.,

Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,

Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission

Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jun 26, 2002 this sequence version replaced gi:21536031.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

QY 409 ArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSer 428
 DB 23282 AGGCGTAGATCCATTCAAGACTTGACTGTTACTGGGACAGAACCTGGTCAGGTCAGAGT 23341
 QY 429 ArgSerSerProSerValArgMetIleThrThrSerGlyProThrSerGluLysPro 448
 DB 23342 CGGTCATCCAGCCCTAGAGTGGAAATGATGATCTCTCAGACCAACCTCAGAGAACCA 23401
 QY 449 ThrArgSerHisProTirThrProAspSerThrAspThrAsnGlySerAspAsnSer 468
 DB 23402 GCTCGAAGTCACCGTGGAAACCTGATGATCCACAGATACCAATGGCTTAGATACTCC 23461
 QY 469 IleProMetAlaTyThrLeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsn 488
 DB 23462 ATCCCAATGGCATATCTTACACTGGATCACCAGCTACAGCCCTCTAGCGCCGTGCCAAAC 23521
 QY 489 SerLysGluSerMetAlaValPheGluGlnHisCysLysMetAlaGlnGluTyMetLys 508
 DB 23522 TCCAAATTAATTCATGGCAGTGTGTTGAACCAATGTTAAATGGCACAGGATATATGAAA 23581
 QY 509 ValGlnThrGluIleAlaLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAsp 528
 DB 23582 GTTCAACCGAAATCGCATTTGTTACTACAGAGAAAGCAGAGTAGTTGCAGAAATTGGAC 23641
 QY 529 GlnAspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeu 548
 DB 23642 CAGGATGAAAGGACACGACAAATATACATCTCTGTTGATACAGGAAACATAAAAGCTTTTA 23701
 QY 549 AspGluAsnLysSerLeuSerThrTyThrGlnGlnCysLysLysGlnLeuGluValIle 568
 DB 23702 GATGAAACAAAGGCTTTCTTACTATTATACAGCAATGCAAAACAACTAGAGGTTATC 23761
 QY 569 ArgSerGlnGlnLysArgGlnGlyThrSer 579
 DB 23762 AGAGCCCAACGACGAGAAACGACAGGCACTTCA 23794
 RESULT 13
 XLU92030 2812 bp mRNA linear VRT 18-APR-1998
 LOCUS Xenopus laevis TGF-beta-activated kinase TAK1 mRNA, complete cds.
 DEFINITION U92030
 ACCESSION U92030
 VERSION U92030.1 GI:3057035
 KEYWORDS
 SOURCE
 ORGANISM
 Xenopus laevis.
 Xenopus laevis.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutelesostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 2812)
 Shibuya, H., Iwata, H., Masuyama, N., Gotoh, Y., Yamaguchi, K., Irie, K.,
 Matsumoto, K., Nishida, E. and Ueno, N.
 Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
 EMBO J. 17 (4), 1019-1028 (1998)
 98130593
 MEDLINE
 PUBMED
 9463380
 REFERENCE
 2 (bases 1 to 2812)
 Shibuya, H.
 Direct Submission
 Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
 Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
 JOURNAL
 FEATURES
 Location/Qualifiers
 1..2812
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 341..2191
 /notes="TGF-beta-activated kinase"
 /codon_start=1
 /product="TAK1"
 /protein_id="AAC14008.1"
 /db_xref="GI:3057035"
 /translations="MSATSAEMIETPPVLNFEIDYKEIEVEVWGRGTGVVCKAKW
 RGKDVAIKQIESESRKAFIVELRQLSRVNHPIVKLYGACLNPLVMEYAGGSLY

NVLGAELPYTAAAHAMSWCLQCAQGVAYLHMKPKALIHRLKPPNLLVAGGTVL
 KICFGTACDIQHTMNNKSGAAWPEVFGSNYSKCDVFSWGIILWEIVTRKPF
 DEIGPAPRIWAVHNGTRPPLIKNLKPIESLMTKWSKDPORPSMWEIVKIMTHL
 MOYEPGADVSLQYPCQYSDGQSNASSTGSCIDITSTNTSNKSDIHIEPDPFOASAS
 NDKTRIESKLAQHLKQAKOTGESRLUSLPPSGSSVESUSEIRGRPPSTLTGSEK
 RMDAMSLEARIISASTAYVKPRGRKRTASFGNILDVPEIITAGNGQRRRSVQDL
 TIVGTESQESRNSRSSPSVRMTTSGPTDPKPRGLPWAPDESDTSDNSIIM
 AVLTLHQLOPLAPCNSKESMAVFEQHKMAQYMKVQTEIALLLQRKOBLIAELDQ
 DEKQONTSLRVQEHKLLDENKSLSTVYQCKKQLEVRISQQQKRGQTS"
 BASE COUNT 808 a 612 c 669 g 723 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,75e-154 Length: 2812
 Score: 2638.50 Matches: 519
 Percent Similarity: 87.50% Conservative: 20
 Best Local Similarity: 84.25% Mismatches: 28
 Query Match: 87.54% Indels: 49
 DB: 5 Gaps: 7
 US-09-830-144-2 (1-579) x XLU92030 (1-2812)
 QY 12 SerSerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGluGlu 31
 DB 344 TCTGCCACCTCTGCTGAATGATTGAGACTCCA---CCGGTCTGTAACCTTTGAGGAGATC 400
 QY 32 AspTyLysGluIleGluValGluValValGlyArgGlyAlaPheGlyValValCys 51
 DB 401 GACTACAGGAGATTGAGGTGGAAGAGGTGTTGGAAGAGGAACCTTTGGTGTCTGTCGC 460
 QY 52 LysAlaLysTrpArgAlaLysAspValAlaLysGlnIleLysGlnLeuSerGluSerGluArg 71
 DB 461 AAAGCTAATGCGGAGGAAAGATGTGGCAATCAACAAATTTGAAAGCGAATCTGAAGG 520
 QY 72 LysAlaPheIleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLys 91
 DB 521 AAAGCCTTTATTGTTGAGCTGCGACAGTGTCTCGTGTTAACCATCCCAATATTGTCAAG 580
 QY 92 LeuTyGlyAlaCysLeuAsnProValCysLeuValMetGluTyAlaGluGlyGlySer 111
 DB 581 TTGTATGGAGCCTGCTAAATCCGGTATGTTTAGTATGGAGTATGCAGAGAGGAGGTTC 640
 QY 112 LeuTyAsnValLeuHisGlyAlaGluProLeuProTyTrpThrAlaAlaHisAlaMet 131
 DB 641 TTGTACAAATGTTTGCATGGAGCTCAACCGCTGCCCTTACTATACAGCTGCTCATCCAATG 700
 QY 132 SerTrpCysLeuGlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAla 151
 DB 701 AGTTGGTGTTTACAATGTGCCAAGAGTTGCATATTTACATAGCATGAAGCGGAAGGT 760
 QY 152 LeuIleHisArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeu 171
 DB 761 CTAATTTACAGGAGCCTCAAAACCACTTTGCTGGTGTAGCTGGAGGACCCCTTCTT 820
 QY 172 LysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGly 191
 DB 821 AAGATTGTGACTTTTGGTACAGCTGTGATTTACAGACTCAGATGACTAATAACAAAGA 880
 QY 192 SerAlaLysTrpMetAlaProGluValPheGluGlySerAsnTySerGluLysCysAsp 211
 DB 881 AGTGACGATGATGGCTCCAGAGTCTTGAAGGTGACCAATTTACAGCGAGAAATGTGAC 940
 QY 212 ValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGlu 231
 DB 941 GTATTAGTTGGGAATAATTTCTTTGGGAAGTAATAACCGAAGAAAACCTTTTCATGAA 1000
 QY 232 IleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeu 251
 DB 1001 ATTGGTGTCCAGCGTTCGGTATCATGTGGCTGTTCATATGGTACTCGCCCACTTA 1060
 QY 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
 DB 1061 ATTAAAAATTTGCCTTAAGCCTATTGAAAGCTTAATGACTCGTGTGTTCCAAAGATCCC 1120

```

QY 272 SerGlnArgProSerMetGluGluIleValIleMetThrHisLeuMetArgTyrPhe 291
    |||||
Db 1121 CCACAAAGACCTTCATGAGGAGATGTCAGATATGACTCATCTAATGACGATATTT 1180
QY 292 ProGlyValAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyIleSerAsn 311
    |||||
Db 1181 CCTGGAGCGGATGTTCTTACAGTACTCTTCAGTACTCAGATGAAAGGCGAAAGTAT 1240
QY 312 SerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSer 331
    |||||
Db 1241 TCTGCCACAAGTACAGCTTCCTGCATGATATTAATCACTAATACTAGTAAACAAGGT 1300
QY 332 AspThrAsnMetGlu-----GlnValProAlaThrAsnAspThrIleLysArg 347
    |||||
Db 1301 GACATCCATATAGAACCTGGTATTTCCAAAGCAAGTCGAGAAATGACAACTCAACGCA 1360
QY 348 LeuGluSerLysLeu-----LeuLysAsnGlnAlaLysGlnGlnSerGluSerGly 364
    |||||
Db 1361 ATTGAATCAAAAGCTTCACACAACCTGAAAAATCAAGCAACAAACGGGTGAATCGGCT 1420
QY 365 ArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeu----- 379
    |||||
Db 1421 CGATTAGCTTGGCTCCCTCTCGGGGTAGCAGTGTAGACAGCTGTCTGAATACGAGCT 1480
QY 380 ---ProPro-----ThrSerGluGlyLysArgMetSerAlaAspMetSerGlu 394
    |||||
Db 1481 AGACCAACCATCTACTCTGGGTACATCTGAGGGGAAACGAATAGTGCAGACATGCTGAA 1540
QY 395 IleGluAlaArgIleAlaAlaThrThr----- 403
    |||||
Db 1541 TTGAAAGCAAGAAATTTCTGCAAGTACAGCATATGTCAAGCTTAAACGAGCCAGTAAA 1600
QY 404 -----GlyAsnGly 406
    |||||
Db 1601 ACTGCTTCATTTGGCAACATTCGATGTCGCCAGATATCATCTTACAGCAGGAAACGGG 1660
QY 407 GlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnVal 426
    |||||
Db 1661 CACACAGAGGGCAGATCTGTACAAAGACCTCACTATGTTGAAACAGAGTCCAGTCAGAG 1720
QY 427 -----SerSerArgSerSerSerProSerValArgMetIleThrThrSerGlyPro 443
    |||||
Db 1721 AGCAGAAACAGTACAGATCTGTACTGCAAGGTAGAGATGATTACTACCTCAGACACA 1780
QY 444 ThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThrAsn 463
    |||||
Db 1781 ACCCCAGACAAAGCCCCACGCTGCTCCATGAGGCCGCCCGATCAATCTTCAGATACAAAT 1840
QY 464 GlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnProLeu 483
    |||||
Db 1841 GGAATCGATTAACCTCACTACTGATGACATGCTTACACTGATCATCAGTTACGCTCTTA 1900
QY 484 AlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCysLysMetAla 503
    |||||
Db 1901 GCACCGGTCTCAAACTTAAGAAATCTATGGCTGTTTGAACAAACACGCAAAATGGCA 1960
QY 504 GlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLysGlnLeu 523
    |||||
Db 1961 CAGAGAGTACATGAAGTGCAGACCGAAATGCTTTATTGCTGAAAGAAAGCAAGAGCTT 2020
QY 524 ValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGlu 543
    |||||
Db 2021 ATTGCAGAACTAGATCAGATGAAAAAAGACACACAGAACATCGGCTCTGGGTCAAGAA 2080
QY 544 HisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGlnCysLysLys 563
    |||||
Db 2081 CACAAAAAGCTTTAGATGAAAAATAAAGCTTTCCACATATATACGCAATGTAAAAAA 2140
QY 564 GlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
    |||||
Db 2141 CAGTTGAAGTCAATCGAAGTCAGACGACAGAAACGACAGAACTTCC 2188

```

```

RESULT 14
HSM800550      3482 bp      mRNA      linear      PRI 18-FEB-2000
LOCUS
DEFINITION
Homo sapiens mRNA; cDNA DKFZps86F0420 (from clone DKFZps86F0420);
partial cds.
ACCESSION
AL050393
VERSION
AL050393.1
KEYWORDS
GI:4914614
SOURCE
Homo sapiens.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 3482)
AUTHORS
Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE
Direct Submision
JOURNAL
Submitted (15-MAY-1999) MIPS, Am Kiofexerplatz 18a, D-82152
Martinsried, GERMANY
COMMENT
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;
sequenced by AGOHA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZps86F0420) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
source
1..3482
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZps86F0420"
/tissue_type="uterus"
/clone_lib="586 (synonym: hute1). Vector pSport1, host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1729..2511
/gene="DKFZps86F0420"
1729..2511
/gene="DKFZps86F0420"
/note="TGF-beta activated kinase 1a, unspliced"
CDS
/gene="DKFZps86F0420"
1729..2511
/translation="MDIANTNTSNKSDTMEQVPAITNDTIKRLSKLKNQAKQSES
GRSLGASGSVSESLPTSEGRKMSADSEIEARLAATGNGGPRRSIQDLVTYGT
EPGVASRSRSPVMTTTSPTSEKPTSHPTPDSDTNGSDRSIMAYLVLTDHO
LOPLACPNKESMAVFEQHKMAQBYMVQTEILLLQKQELVAELDDQDKDQNT
SRVQEHKLLDKNKSLSTYYQCKQKQLEVINQQKRGQTS"
polya_signal
3297..3302
polya_site
3326
BASE COUNT
1222 a 568 c 620 g 1072 t
ORIGIN
Alignment Scores:
Pred. No.: 3,04e-82 Length: 3482
Score: 1476.00 Matches: 289
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 1
Query Match: 48.97% Indels: 0
Gaps: 0
DB: 9
US-09-830-144-2 (1-579) x HSM800550 (1-3482)
QY 290 TyrPheProGlyValAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGln 309
    |||||
Db 1639 TACTTTCAGAGCAGAGATGAGCCATTACAGATCTTGTGACGATTCAGATGAGAGCAG 1698
QY 310 SerAsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsn 329
    |||||
Db 1699 AGCAACTGCGCACACAGTACAGGCTCATGACATTCGCTTCAATAATACAGATTAAC 1758
    |||||
QY 330 LysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlu 349
    |||||

```

Db	1759	AAAAAGTGACACTAATATGGAGCAAGTTCCTGCCACAATGATCTATTAAAGCGCTTAGAA	1818	
Qy	350	SerLysLeuLeuLysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeuGly	369	
Db	1819	TCAAAAATGTTGAAAAATCAGGCACACAGAGTGAATCTGGAGCTTTAAAGCTTTGGGA	1878	
Qy	370	AlaSerHisGlySerSerValGluSerLeuProProThrSerGluGlyLysArgMetSer	389	
Db	1879	GCCTCCGTCGGAGCAGTGTGAGAGCTTGCCTCCCAACCTCTCAGGGCAAGAGATGAGT	1938	
Qy	390	AlaAspMetSerGluLeuGluAlaArgIleAlaAlaThrThrGlyAsnGlyGlnProArg	409	
Db	1939	GCTGACATGCTCTGAAATAGAGCTAGGATCGCGCAACACAGCCCAACGACCAAGA	1998	
Qy	410	ArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArg	429	
Db	1999	CGTAGATCCATCCAGACCTTGACTGTAACCTGGACAGACCTGGTCAGGTGACGATAG	2058	
Qy	430	SerSerSerProSerValArgMetIleThrThrSerGlyProThrSerGluLysProThr	449	
Db	2059	TCATCAGTCCCACTGTGTGAGATGATTAACCTCAGGACCAACCTCAGAAAAGCCAACT	2118	
Qy	450	ArgSerHisProThrThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIle	469	
Db	2119	CGAAGTCATCCATGGACCCCTGATGATTCACAGATACCAATGGATCAGATCACTCCATC	2178	
Qy	470	ProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSer	489	
Db	2179	CCAATGGCTTATCTTACCTGGATCACCACCTACAGCCTTAGCACCGTCCCAAACTCC	2238	
Qy	490	LysGluSerMetAlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysVal	509	
Db	2239	AAAGAATCTATGGCAGTGTGTGAACAGCATGTGAAAATGGCAAGATATATGAAAGTT	2298	
Qy	510	GlnThrGluIleAlaLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGln	529	
Db	2299	CAACAGAAATTCATGTTTATTACAGAAAGCAAGAACTAGTTCGAACTGGACCAG	2358	
Qy	530	AspGluLysAspGlnGlnAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAsp	549	
Db	2359	GATGAAAAGGACCAGCAAAATACATCTCGCTGGTACAGAAACATATAAAGCTTTAGAT	2418	
Qy	550	GluAsnLysSerLeuSerThrTyrGlnGlnCysLysLysGlnLeuGluValIleArg	569	
Db	2419	GAAAACAAAAGCCTTTCTACTTACTACCAAGCAATGCAAAAAACAACAGAGTCATCAGA	2478	
Qy	570	SerGlnGlnGlnLysArgGlnGlyThrSer	579	
Db	2479	AGTCAGCAGCAGAAACACAGGCATTTCA	2508	
RESULT 15				
AF199466				
LOCUS				
DEFINITION Drosophila melanogaster TGF-beta activated-kinase 1 homolog mRNA,				
complete cds.				
ACCESSION AF199466				
VERSION AF199466.1				
KEYWORDS				
SOURCE				
ORGANISM				
Drosophila melanogaster.				
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
Ephydroidea; Drosophilidae; Drosophila.				
1 (bases 1 to 3349)				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
2 (bases 1 to 3349)				

AUTHORS

TITLE

JOURNAL

FEATURES

source

CDS

Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M., Matsumoto, M., O'Connor, M. B., Shibuya, H. and Ueno, N.

Direct Submission

Submitted (28-OCT-1999) Developmental Biology, National Institute for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi 444-8585, Japan

Location/Qualifiers

1..3349

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/chromosome="X"

/map="19E5-7"

/19E5..2952

/note="TAK1 homolog"

/codon_start=1

/product="TGF-beta activated-kinase 1 homolog"

/protein_id="AAF06815.1"

/db_xref="GI:6289099"

/translation="MATASLDALQAAVDFSEITLREKVGHSYGVVCAKAVWRDKLVA VKFFASAEOKDIEKEVKQLSRVHPNIIALHGISSVQOATYLI MFASGGSLHNEFLH GKVKPAYSLAHAMSWARQCAEGLAYLHMTPKPLIHRDVKPLNLLITNKRNLKICDF GTVADKSTMMNNRGAAMWAEVFEQSKTEKCDIPFSAIWLVEVLSRKQPPKPGIDN AVTIQWKIKYGERPPLITTCPKRIEDLMTACWKTPEDRPSMQYIVGVNHEIVKDYTG ADKALYTFVYVQIVTKESDGTVAQPDLSLQSEGLSPSTQITPTTAANANVAIA ISKTTTSSMTENTSSDITPTNSGOLDNNPLFYVWVTRWDALPEESNESRNDSPN LTSABATORLETIRNGMILMACKPWEQLTLDVEANGFDLSPSESSSTNKAISDRE RLVTDTKPVMTTDLNNNGGIHAHNSGILLSHANGWAQARDBELQBEHQEIVNSDL VYDVPDENDGTQEQLAELDFELQPEPFPNDASQLIYRDRHMAKEYIUSVDTNL YVADPKDLIVQMDREHQELLRKMKEGLQSLYNNLQOQYASRLAAGHHHPQ PHPHPHNQLOHPHSPHMFHFDQEGCLLPGSVCGGSESVEGWWVIPPHNNA"

BASE COUNT

934 a 893 c 853 g 669 t

ORIGIN

Alignment Scores:

Pred. No.: 5.15e-51

Score: 971.00

Length: 3349

Matches: 232

Percent Similarity: 52.22%

Conservative: 97

Best Local Similarity: 36.83%

Mismatches: 202

Query Match: 32.22%

Indels: 100

DB: 3

Gaps: 14

US-09-830-144-2 (1-579) x AF199466 (1-3349)

Qy

17

GluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIle

36

|||||

Db

913

GAGATGCCACAGCATCGCTGGACGCACTCCAGGCAGCCTATGTGGACTTCAGTGAGATA

972

|||||

Qy

37

GluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArg

56

|||

Db

973

ACACTAAGAGAGAAAGTCGGCCATGGGTCTCTCGTGTGGTCTGCAAGGCCGTTTGGCGC

1032

|||

Qy

57

AlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal

76

|||

Db

1033

GACAGCTGGTTCGGTCAAGGAGTTCCTCCAGCGCCGACGAGCAAGGACATCGAGAAG

1092

|||

Qy

77

GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly-----

94

|||

Db

1093

GAGGTGAAGCAGTTGTGCGCGTGAAGCACCACCAATATCGCTCTGCACGGATATCC

1152

|||

Qy

95

AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn

114

|||

Db

1153

TCGTACCAGCAGGCCACCTACCTGATAATGGATTCCGCGAAGGTGGATCGCTGCACAC

1212

|||

Qy

115

ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys

134

|||

Db

1213

TTCTCTCACCGC---AAGGTGAAGCCCGGCGATATCTCTGCGCCACCGCATGAGCTGGCG

1269

|||

Qy

135

LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis

154

|||

Db

1270

CGCCAATGTCAGAGGGTCTGGCATATTTGCATGCCATGACGCCAAAACCACTAATACAT

1329

|||

Qy

155

ArgAspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCys

174

|||

Db 1330 CGCGACGTGAAGCCGCTGAACCTGCTTGTGACCAACAAAGGACGCAATCTGAAGATATCC 1389
 QY 175 AASPHEGLYTHRALACYSASPILLEGINTHRHISMETTHRASNANSLYSGLYSERALAALA 194
 Db 1390 GACTTCGGCAGCGGTGGCGGACAAATCGACATGATGACCAACAAATCGGATGATGGCGCT 1449
 QY 195 TRPMETALAPROGLIVALPHEGLUGLYSERASNLYRSERGLULYSYSPVALPHESER 214
 Db 1450 TGGATGGCGCCCGAGAGTCTTCGAAGGCTTCAAGTATACGGAAAGTGTGACATTTTCAGC 1509
 QY 215 TRPGYLIEILEUTRGLIVALLIETHARGARGLYSPROPHASPGULIEGLYGLY 234
 Db 1510 TGGGCAATGTTCTATGGAGGTTCTGTCCAGAGAGCGCCCTTTAAAGCATGACAAAT 1569
 QY 235 PROALAPHEARGILEMETRPAALVALHISANGLYTHRARGPROPROLEULELAYSASN 254
 Db 1570 ---GCCCTACACCATCCAGTGAAGATCTACAAAGGTGAACGCCGCCGCTGTGACCACT 1626
 QY 255 LEUPROLYSPROILEGLUSERLEUWETTHARGYSTIPSERLYASPPROSERGLINARG 274
 Db 1627 TGCCCAAGCGCATCGAGGAGCTGATGACCGCTGTGAAAAACGGTGCAGGATGCG 1686
 QY 275 PROSERMETGLUGLULIEVALYSLIEMETTHRHISLEUMETARGTYRPHEROGLYALA 294
 Db 1687 CGGTGATGACATGACATGATGGGCGTTATGACAGAGATGTCAGAGACTATACGGGGCG 1746
 QY 295 ASPGLUPROLEUGLINTYR----- 300
 Db 1747 GACCAAGCCCTGGAATACATTTGTTATCAACAGATTGTCCACCAAGAGAGCGACGC 1806
 QY 301 -----PROCYSGINTYRSEASPGLUGLYGLINSERASN----- 311
 Db 1807 ACGGTGGCCGCTCAACCGGATAGCTCAAGTTCGACAGAGGGGAACTGAGCCCTGCTCC 1866
 QY 312 -----SERALA 313
 Db 1867 ACACAGTTAACCGCAACGCGGCCCAACGCCAATGTGAACGGATAGCAATATCAAAA 1926
 QY 314 THRSETHRGLYSERPHEMETASPILEALASERTHRASNTHRSERASNLYSERASPTHR 333
 Db 1927 ACACAGACTAGCTCATGACCGAAATACCTCATCAACATCATCGGACATCACGCCGACG 1986
 QY 334 ASMETGLUGLINVALPRO-----ALATHRASN-----ASPTHR 344
 Db 1987 AACTCGGCGCACTGACCAATATCCGTAATTCATATATGATGATCCCAATGCTGGGACGCG 2046
 QY 345 ILELYARGLEUGLUSERLYSLEULELAYSASN----- 355
 Db 2047 ATTCCCGAGAGAGAGCAACGAGAGCCGGAACGATAGCTTCAACCTCACTCGTCGGCT 2106
 QY 356 GLNALALYSGLINSEARGLEUGLYARGLEUSERLEUGLYALASERHISGLYSER 375
 Db 2107 GAGGCCACTGAGCGCTCGAAACGATCCGGAACGCGCATGATCCTGATGCGCTGCAAGCCC 2166
 QY 376 VALGLUSERLEUPROTHRSERGLUGLYLYARGMETSERALASPMETSERGLULE 395
 Db 2167 ATGAGAGAGCTC-----ACCTCGACGTGAGCGGAATGGCTTTGATCTGAGTCCAGC 2220
 QY 396 GLUALAARGILEALALATHRTHRGLYASNGLYGLINPROARGARGSERILEGINASP 415
 Db 2221 GAADAGCAGCGACGACGCAACGCAAGAGGATGGCCGCAACGA----- 2268
 QY 416 LEUTHYVALTHRGLYTHRGJUPRO----- 423
 Db 2269 CTCACGCTGACGGACCAACGCGGTGATGATGACCAAGGATCTGTCCAACAACAGCGC 2328
 QY 424 GLY-GLINVALSERSERASERSETPROSERVALARGMETILETHRTHRSERGLYPR 443
 Db 2329 GGCATCCACCGCCCACTGAAAGGACTGCTGAGGCAATGGTT-----GGCAA 2379
 QY 443 CTNRSERGLULYSPROTHRARGSERHISPRO-----TRPTh 455
 Db 2380 GCAAGAGATGAGAGCTGCGAGAGCAAGACATGAGCAGAGATTTGTCAACTCGTTGGAC 2439

QY 455 RPROASPASERTHRASPTHRASNGLYSERASPSANSERIIEPROMETALATYRLEUTH 475
 Db 2440 GTC-GACGTGGATCCGACGAGGATGAGAACGACGACCAACAGTCACTGCGGAGAT 2498
 QY 475 RLEUASPHISGLINLEUGLIPROLEUALPROCYSPROANSERTLYSGIUSERMETALA 495
 Db 2499 TCTTGATCCGGAAGCTCCAGCCAGAGCCCGATATCCCAACGATGCCGATGCCACTCAT 2558
 QY 495 IPHEGLUGLNIHISCYSLYMETALAGLINTYRMETLYVALGLINTHRGLUILEALE 515
 Db 2559 CTACCGGACACACGACATGGCCCAAGGATGACCTGACGCTGACACGACCTTACTRA 2618
 QY 515 ULEULEGINARGLYGLINLEUVALAGIULEUASPGLINASPGLULYASPGINGL 535
 Db 2619 CGCGCAGGACTTCMAAGACAGCTCATCTGATGATGAGCCGAACCGAGCGCAACAGAA 2678
 QY 535 NANTHRSERARGLEUVALGLINLHISLYSLYSLEULEUASPGLUASNLYSERLEUSE 555
 Db 2679 G-----CAGGAGCTTCTCGCGCAAGATGAAGACAAAGAGAGGTCTTCA 2720
 QY 555 RTHRTRYRGLINGLINCYSLSYSLGLN 564
 Db 2721 GAGCTTTACAAACATCTGACGACGAG 2748

Search completed: December 10, 2002, 01:08:33
 Job time : 3724.32 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:57:35 ; Search time 286.56 Seconds
(without alignments)
4550.213 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTASAASSSSSSAGEMIE.....QCKQLEVIRSQOQKRGTS 579

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/sgn2_1/USPTO.spool/US09830144/runat_04122002_141353_2242/app_query.fasta_1.1422
-DB=N_Geneseq_101002 -QFMT=fastp -SURFIX=p2n.rng -MINMATCH=0.1 -LOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144.ecgn 1 113 @runat_04122002_141353_2242 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MWAP -LARGQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq_101002.*

1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	3014	100.0	1788	20	AA56285	Human TAK1-6xHis e
2	3014	100.0	1959	18	AAT85095	Human transforming
3	3014	100.0	2656	20	AA56279	Human TAK1 encodin
4	3014	100.0	2656	21	AAA39105	Human TAK-1 nucleo
5	3006	99.7	2769	24	ABL88437	Pain regulated cDN
6	3006	99.7	2785	20	AA59696	Human TGF-beta act
7	2986	99.1	2443	18	AAT85094	Mouse transforming
8	2982.5	99.0	2866	20	AA59697	Human TGF-beta act
9	2746	91.1	1704	20	AA59698	Human TGF-beta act
10	971	32.2	3367	23	ABL02489	Drosophila melanog
11	702.5	23.3	10997	23	ABL02488	Drosophila melanog
12	505	16.8	759	23	ABL08337	Drosophila melanog
13	493.5	16.4	3454	24	ABL70018	Pancreas cancer re
14	490	16.3	2120	21	AA599726	cDNA encoding huma
15	482	16.0	1706	21	AAA75674	DNA encoding a hum
16	482	16.0	2191	22	AAH99263	Human protein enco
17	482	16.0	2194	21	AA599734	Cardiovascular sys
18	482	16.0	2220	21	AA599734	Human survival reg
19	481	16.0	1365	21	AA599727	cDNA encoding huma
20	481	16.0	2272	21	AA599736	Cardiovascular sys
21	469	15.6	3141	22	AAD18824	Human kinase (PKIN
22	469	15.6	3538	24	AAD34309	Human PKIN-12 cDNA
23	462	15.3	3066	24	ABO86165	Novel human gene.
24	462	15.3	3558	24	ABK83874	Human cDNA differe
25	462	15.3	3558	24	AA361339	Human mitogen acti
26	461.5	15.3	2157	22	AAH46913	cDNA encoding huma
27	461.5	15.3	3111	24	ABN86357	Novel human protei
28	461.5	15.3	3518	24	ABN86358	Novel human protei
29	457.5	15.2	2622	22	AA575336	Human TGF-beta rec
30	457.5	15.2	3967	22	AAH73366	Human cervical can
31	456.5	15.1	2403	22	AAF44701	Novel protein kina
32	436.5	14.5	3072	23	ABL04365	Drosophila melanog
33	434.5	14.4	1224	21	AAC43254	Arabidopsis thalia
34	434	14.4	1063	22	AAH34976	Human colon cancer
35	434	14.4	1631	21	AAC39537	Arabidopsis thalia
36	427.5	14.2	1591	21	AAC40839	Arabidopsis thalia
37	426.5	14.2	2283	21	AAC48526	Cardiovascular sys
38	425	14.1	2254	21	AA599735	Drosophila melanog
39	424	14.1	3109	23	ABL29755	Human CARK (Cardia
40	418.5	13.9	3025	21	AAA47606	Cardiovascular sys
41	418	13.9	2069	21	AA599737	Human CARK (Cardia
42	418	13.9	2505	21	AAA47607	Novel protein kina
43	418	13.9	2508	22	AAF44702	DNA encoding novel
44	418	13.9	2788	23	AA580720	Rat CARK (Cardiac
45	418	13.9	3026	21	AAA47608	

ALIGNMENTS

RESULT 1
AA56285
ID AA56285 standard; DNA; 1788 Bp.
XX
AC AA56285;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAK1-6xHis encoding DNA.
XX
KW Human; TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
OS Homo sapiens.
OS Synthetic.
XX
XX Key
FH Location/Qualifiers
FT 7..1779


```

FT      /cag= a
XX      MO9921010-A1.
PN      29-APR-1999.
XX      22-OCT-1998; 98MO-JP04796.
XX      22-OCT-1997; 97JP-0290188.
XX      (CHUS ) CHUGAI SEIYAKU KK.
PI      Ohtomo T, Ono K, Tsuchiya M;
XX      WPI: 1999-312645/26.
DR      P-PSDB: AAY09547.
XX      Screening for TGF- beta inhibitory substances, which are useful as
PT      drugs for treatment of diseases relating to its disorder
XX      Example 1; Page 167-171; 195pp; Japanese.
XX      A method has been developed for screening for substances which inhibit
CC      the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC      comprises: (a) contacting the polypeptide in the presence of a sample;
CC      and (b) detecting the amount of bound polypeptide, in which the sample
CC      can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC      growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC      indications e.g. as TGF-beta signal transduction inhibitors or
CC      activators, or extracellular matrix protein production enhancement
CC      inhibitors, or activators, or cell proliferation prevention inhibitors or
CC      activators, or monocyte migration inhibitors or activators, or
CC      physiological activity induction inhibitors or activators, or
CC      immunosuppression inhibitors or activators, or amyloid beta protein
CC      precipitation inhibitors or activators, and such substances can also be
CC      inhibitors of the TAK1 polypeptide function, particularly kinase
CC      activity. The present sequence encodes TAK1-6xhis from an example of
CC      the present invention.
XX      Sequence 1788 BP: 551 A; 394 C; 413 G; 430 T; 0 other;
SO
Alignment Scores:
Pred. No.: 3,296-217 Length: 1788
Score: 3014.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-2 (1-579) x AAY56285 (1-1788)
OY      1 MetSerThrAlaSerAlaIaIaSerSerSerSerSerAlaGlyGluMetIleGlu 20
DB      7 AITCTACAGCCCTCTGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 66
OY      21 AlaProSerGlnValIleuasnPhenGluGluIleAspTyrLysGluIleGluValGluGlu 40
DB      67 GCCCTTCCAGGTCCTCAACTTGAAGAGATCGACTTAAGAGAGATCGAGGTGAGAG 126
OY      41 ValValGlyArgGlyAlaPhenGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
DB      127 GTTGTGGAGAGAGAGCTTTGAGAGTTGTTGCAAGAGCTTAAGTGAAGACCAAGATGTT 186
OY      61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB      187 GCATTTAACAAATAGAAAGTGAATCTGAGAGAAAGCCGTTTATGTAGAGCTTCGCGAG 246
OY      81 LeuSerArgValIasnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
DB      247 TTATCCGCTGTAACCATCTTAATATGTAAAGCTTTAAGAGACCTGCTGAATCCAGTG 306
OY      101 CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValIleuHisGlyAlaGlu 120

```

```

DB      307 TGTCTTGATGATGAATATGCTGAAGGGGCTCTTATATATATGTCGATGTCCTGA 366
OY      121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB      367 CCATTGCCATATTATATATGCTGCTGCCACGCAATGAGTGTGTGTTACATGTTCCCAAGA 426
OY      141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAspAspLeuLysProPro 160
DB      427 GTGGCTTATCTTCACAGCATGSCAACCCAGGCTTAATTCACAGGAGACCTGAACACCA 486
OY      161 AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB      487 AACTTACTGCTGTTGACAGGGGGGACAGTTCTAAATAATTGTTGATTGTACACGCTGT 546
OY      181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
DB      547 GACATTCAAGACACATGACCAATTAACAAGGAGATGCTGCTGATGGCACCTGAAGTT 606
OY      201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB      607 TTGAAGGTATGTAATTACAGTGAATAATGTACGCTTCAGCTGGGGTATTATCTTTGG 666
OY      221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
DB      667 GAAGTATACCGGTCGGAACCCCTTGTATGAGATTGGTGCCCGCTTCGAAATCATG 726
OY      241 TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB      727 TGGCGTTCATATATGTTACTCGACACCATGATMAAAATTTACTTAAAGCCCATGAG 786
OY      261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
DB      787 AGCTGATGACTGTTGTTGCTTAAGATCCTTCCAGCGCCCTTCATGAGGAAAT 846
OY      281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
DB      847 GTGAAATATATGACTCATCTGATGCGGTACTTCCAGGAGCATAGCATTAACGTAAT 906
OY      301 ProCysGlnTyrSerAspGluGluGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
DB      907 CTTGTGATTCAGTTCAGTGAAGGACAGAGCACTTCGACACGTAACAGGCTCATTCATG 966
OY      321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
DB      967 GACATTCCTCTCAATAATAGAGTAACAAGAGCACTAATATGAGCAAGTCTCTGCC 1026
OY      341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
DB      1027 ACAATGATATCTATTAAAGCCCTTAAGATCAAAATTTGTGAATAATCAGGCAAGCAAG 1086
OY      361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
DB      1087 AGTAATCTGACGTTTAACCTTGGAGCTCCCAATGGAGCGATGTGAGAGCTTGCCC 1146
OY      381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
DB      1147 CCAACCTCTAGGAGGACAGAGAGATGATGCTGATGCTGAATAATGAAGCTAGATCGCC 1206
OY      401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
DB      1207 GCAACCAAGGCAACGAGACGCAAGACGTAGATCCATCCMAAGCTTGAATGTAACCTGGA 1266
OY      421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
DB      1267 ACAGAACCTGTCAGGTGAGACGATGATCATCCAGTCCAGGTGCAAGATTAATCACTACC 1326
OY      441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
DB      1327 TCAGGACCAACCTCAGAAAGCCAACTCGAAGTCATCATCAGAACCCCTGTGATTCACACA 1386
OY      461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrIleuThrLeuAspHisGlnLeu 480
DB      1387 GATACCAATGATGATGATTAATCTCATCCCAATGSCTTATCTTACACTGATCAACCACTA 1446

```



```

OY 321 AsplleAAserThrasnThserAsnlysserAspThrasMerGluGluValProAla 340
DB 1143 GACATTGCTTCTACAAATACGATACAAAGATGACACATATATGAGACAAAGTTCTGCC 1202
OY 341 ThrasAspThrlleYsarGluGluSerlySLeuLeuAlaYsGlnGln 360
DB 1203 ACMAATGATACATTAAGGCTTAGAATCAAAATTGTTGAAAAATCAGCAAGCAAG 1262
OY 361 SerGluSerGlyYarGluSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
DB 1263 AGTGAATCTGACGTTTAAGCTTTGAGAGCTCTCCATGGAGACAGTGTGAGAGCTTGCCC 1322
OY 381 ProThrSerSerGluGlyYsarGmerSerHisAspMetSerGluGluAlaArgIleAla 400
DB 1323 CCACCTCTGAGGGCAAGAGATGAGTGTGACATGTCTGAATAGAAAGCTTGATCGCC 1382
OY 401 AlaThrThrGlyYAsnGlyValnProArGArGArSerIleGlnAspLeuThrValThrGly 420
DB 1383 GCACACACAGGCAACGACAGCCAGACAGTACATCCATCCAAAGACTTGACTGTACTGGA 1442
OY 421 ThrGluProGlyValnYsarSerArGSerSerSerProSerValArgMetIleThrThr 440
DB 1443 ACAGAACTGTGTAGGTGACAGTGTGATCATCCAGTCCAGTGTGAGATGATTTCTACC 1502
OY 441 SerGlyProThrSerGluYsarProThrArGSerHisProTrpThrProAspSerThr 460
DB 1503 TCAGAGCCAACTCAGAAAAAGCCAACTCAGAGTCAATCCATGAGACCCCTGATATTCACA 1562
OY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaYLeuThrLeuAspHisGlnLeu 480
DB 1563 GATACCAATGATCAGATCAATACCTCAATCCCAATGGCTTATCTTAACCTGATCACAAC 1622
OY 481 GlnProLeuAlaProCysProAsnSerlySLeuSerMetAlaValPheGluGlnHisCys 500
DB 1623 CAGCCTCTGACAGCCGTGCCCAAACTCCAAAGATCTATGAGTGTTCAGACAGCATTTGT 1682
OY 501 LysMetAlaGlnGluYsarMetlySValGlnThrGluIleAlaLeuLeuGlnArgLys 520
DB 1683 AAAATGGCAACAAATATATGAAAGTTCAACAGAAATGCAATGTTATTACAGAGAAAG 1742
OY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluYsarGlnGlnAsnThrSerArgLeu 540
DB 1743 CAAGAACTAGTTCAGAACTGAGCCAGATGAAAGAACAGACAGCAAAATATCATCTGCC 1802
OY 541 ValGlnGluHislySLeuLeuAspGluAsnLysSerLeuSerThrTrpYrGlnGln 560
DB 1803 GTACAGGAACATATAAAAGCTTTAGATGAAACAAAAGCCTTTCTCTTACCTACAGCAA 1862
OY 561 CysLyslySLeuGlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
DB 1863 TGCAAAAAACAACTAGAGGTCAATCAGAGTCAAGCAGCAAGAACAGCAAGCACTTCA 1919

RESULT 3
AA556279
ID AA556279 standard; DNA; 2656 BP.
XX
XX AA556279;
AC
XX
XX 21-JUL-1999 (first entry)
DT
XX
XX Human TAK1 encoding DNA.
DE
XX
XX Human; TAK1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 183..1922
XX FT /*tag= a
XX
XX MO921010-AL.

```

```

PD 29-APR-1999.
XX
XX 22-OCT-1998: 98WO-JP04796.
XX
XX 22-OCT-1997: 97JP-0290188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
PI Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
DR
DR P-PSDB; AA09542.
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder
XX
XX Example 1; Page 150-154; 195pp; Japanese.
PS
XX
XX A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAK1.
XX
XX
SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;

Alignment Scores:
Pred. No.: 5,47e-217 Length: 2656
Score: 3014.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-09-830-144-2 (1-579) x AA556279 (1-2656)
OY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyMetIleGlu 20
DB 183 ATGTCTACAGCCTCTGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 242
OY 21 AlaProSerGlnValLeuAsnProGluGluIleAspTrpTrpGluGluValGluGlu 40
DB 243 GCCCTTCCCAAGTCCCACTTGAAGAGATCACTACAAAGAGATCGAGTGAAGAG 302
OY 41 ValValGlyArgGlyAlaPheGlyValIleCysLysAlaIleStrpArgAlaIleAspVal 60
DB 303 GTTGTGGAAGAGAGACCTTTGCAATGTTTGCACAAAGTGAAGAGAGCAAAAGATGTT 362
OY 61 AlaIleLysGlnIleGluSerGluSerGluArgGlyAlaPheIleValGluLeuArgGln 80
DB 363 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAACGTTTATTGAGAGCTTCCGAG 422
OY 81 LeuSerArgValaLeuHisProAsnIleValLysLeuTrpGlyAlaCysLeuAsnProVal 100
DB 423 TTATCCCGGTGGAACCAATCCATATATGTAAGCTTTATGAGAGCTGCTGAATCCAGTG 482
OY 101 CysLeuValMetGluYrYrAlaGluGlyGlySerLeuYrAsnValLeuHisGlyAlaGlu 120
DB 483 TGTCTTGATGAGAAATGCTGAAGGGGCTCTTTATATATATGCTGCATGCTGCTGAA 542
OY 121 ProLeuProYrYrYrThraAlaAlaAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
DB 543 CCATTGCCATATTAATAGCTGCCACGCAATAGATGAGTGTGTTACAGTGTCCAAAGGA 602

```

QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 603 GTGGCTTATCTTCACAGATGCAACCCAAAGCGCTAATTACACAGGACCTGAAACCA 662
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGCTGGTTGCAGGGGGACAGTCTTAAATTTGTGATTTTGTGTACAGCCTGT 722
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Db 723 GACATTGAGACACATGACCAATAACACAGGGAGTGCTGCTTGGATGGCACCTGAAGTT 782
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 783 TTTGAAGGTAGTAATTAACAGTGAATAATGTGACGTCTTCAGCTGGGGTATTATTCITTTGG 842
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAGTGATAACGGCTGCGAAACCCCTTTTGATGAGATTGGTGGCCCGACGTTTCCGAATCATG 902
QY 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGCTGTTCTAATAATGTTACTCGACCACCTGATGATAAAATTTACTAAGCCCAITGAG 962
QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCCTGATGACTCGTGTGTTGGTCTAAGATCCCTCCAGCGCCCTTCAATGGAGGAAT 1022
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAAAATAATGACTCACTTGTGTCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
QY 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
Db 1083 CCTTGTTCAGTATTCAGATGAGACAGACGAACTCTGCCACCACTGACAGCTCATTCATG 1142
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
Db 1143 GACATCTCTCTCAATAACGATGAGTAACAAAAGTGACACTAATATGAGCAAGTTCTCTGCC 1202
QY 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1203 ACAATGATACTATTAAAGCCCTTAGAATCAAAATTTGTGAAAAATCAGGCAAGCAACAG 1262
QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
Db 1263 AGTGAATCTGGAGCTTTAAGCTTTGGAGCCCTCCATGGGAGCAGTGTGGAGACTTTGCC 1322
QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1323 CCAACCTCTGAGGGCAAGAGGATGAGTGTGACATGTCTGAAATAGAAGCTAGGATCGCC 1382
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1383 GCAACACAGCGCAACGGACAGCGCAAGCGTAGATCCATCCAAAGACTTGACTGTAACGTGA 1442
QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
Db 1443 ACAGAACCTGCTCAGTGACGATGAGTGTATCCAGTCCCGATGTCAGATGATTTACTACC 1502
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTrpThrProAspAspSerThr 460
Db 1503 TCAGGACCAACCTCAGAAAAAGCAACTCGAAGTCATCCATGGACCCCTGTGATTCACA 1562
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
Db 1563 GATACCAATGGATCAGATACTCCATCCCAATGGCTTATCTTACACTGGATCACCACCTA 1622
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
Db 1623 CAGCCTCTAGCACCGTCCCAAACTCCAAAGAAATCTATGGCAGTGTGTGAACAGCAATTGT 1682

QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys 520
Db 1693 AAAATGGCACAGAATATATCAAGTTCAACACAGAAATTGCATTGTTATTATACAGAAAG 1742
QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
Db 1743 CAAGAACTAGTTGCAGAACTGGACAGGATGAAAAGGACCAAGCAAAATACATCTCGCCTG 1802
QY 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrGlnGln 560
Db 1803 GTACAGGAACATATAAAGCTTTTAGATGAAAAACAAGCCCTTTCTACTTACTACAGCAA 1862
QY 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
Db 1863 TGCAAAAACAACATAGAGTTCATCAGAGTCAGCAGCAGAAACGACAGGCACTTCA 1919

RESULT 4

AAA39105
ID AAA39105 standard; DNA; 2656 BP.

XX
AC AAA39105;

XX
DT 04-SEP-2000 (first entry)

XX
DE Human TAK-1 nucleotide sequence SEQ ID NO:1.

XX
KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
screening; signal transduction; inhibition; inflammatory cytokine;
IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
antinflamatory; suppression; ds.

XX
OS Homo sapiens.

XX
FH Key Location/Qualifiers
FT CDS 183..1922

FT
FT /*tag= a
/product= "TAK-1"

XX
PN WO200023610-A1.

XX
PD 27-APR-2000.

XX
PF 21-OCT-1999; 99WO-JP05817.

XX
PR 21-OCT-1998; 98JP-0299962.

XX
PA (CHUS) CHUGAI SEIYAKU KK.

XX
PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

XX
WPI: 2000-339707/29.

XX
P-PSDB; AAY91000.

XX
PT Method for screening inhibitors of TAK1 signal transduction for
suppression of inflammatory cytokine production and use as
antiinflammatory agents -

XX
PS Example 1; Page 73-80; 100pp; Japanese.

XX
CC The present invention describes a method for screening compounds for
inhibition of inflammatory cytokine signal transduction by contacting
the sample with TAK1 and its receptor TAB1 and selecting for inhibition
of TAK1/TAB1 binding. Also described is a method for screening compounds
for inhibition of inflammatory cytokine signal transduction in which the
inhibition of TAK1 phosphorylation is selected for; and drug
compositions for the treatment of inflammatory disorders containing as
active component an inflammatory cytokine signal transduction inhibitor.
TAK1 is an essential component of the signalling process which results
in release of inflammatory cytokines such as interleukin-1 (IL-1),
IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
for the selection of effective antiinflammatory agents. The present
sequence encodes human TAK-1, which is used in the exemplification of the
present invention.

Db	1663	AAATGCGCAACAAATATATGAAGTTCAAACAAGAAATTGCCTTGTATTACAGAGAAG	1722
Qy	521	GIAIGLNUeVa1aIaGLUeUsApGiNaSpGiULysAsPgiNgInLaSnThSeTaRgLeu	540
Db	1723	CAMGAACtACTGTCAGAACCTGCACCAGCATGAAAAGACACACAAAAATACTATCCGCTG	1782
Qy	541	Va1GIgNLIhiSLySLySLeUeUsAPGiUaAmLySeTeSeTerThTyTYcJIn	560
Db	1783	GTACAGAAACATAAAAAGCTTTAGATGAAAACAAAAGCCTTCTTACTTACACAGCA	1842
Qy	561	CySLySLySGInLUeGIuVallleaYSeRGInGLnLySaRGInGLyThSer	579
Db	1843	TGCAAAAACAACTAAGAGTCATCAGAGTCAGCACAGAAACGAAAGCATTCA	1899
RESULT	6		
AAAY9696	ID	AAAX9696 standard; cDNA to mRNA; 2785 BP.	
XX	XX	AAAX9696:	
AC	DT	18-OCT-1999 (first entry)	
XX	DE		
XX	XX	Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.	
KW	XX	Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;	
KM	XX	TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;	
XX	XX	intractable diseases; atrophic dermatitis; psoriasis; viral infection;	
FT	XX	endotoxin shock; septicemia; human; hTKA1a; ss.	
OS	XX	Homo sapiens.	
FH	XX	Key Location/Qualifiers	
FT	CD5	163..1902	
FT	XX	/tag= a	
FT	XX	/product= "hTKA1a"	
PN	MO9940202-A1.		
XX	12-AUG-1999.		
PD	02-FEB-1999;	99WC-JF00422.	
PF	30-OCT-1998;	98UP-0309316.	
PR	06-FEB-1998;	98JP-0026003.	
XX	(TANABE SEIYAKU CO.		
PA	Hasegawa K, Kageyama N, Sakurai H, Sugita T;		
P1	WPI; 1999-494298/41.		
DR	P-PSDB; AAY28996.		
XX	Nuclear factor kappa B activation inhibitors, useful as preventives		
PT	for, e.g. autoimmune diseases		
PRT			
XX	Examples; Page 35-39; 49pp; Japanese.		
P3	The invention provides a method for identifying or screening a nuclear		
CC	factor kappa B (NF-kB) activating inhibitor by examining the effect of a		
CC	test substance on modulating the function(s) of TGF-beta activated kinase		
CC	1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to		
CC	treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis)		
CC	intractable diseases with inflammation (such as atrophic dermatitis and		
CC	psoriasis), viral infection, endotoxin shock, septicemia and others. The		
CC	present sequence represents the nucleotide sequence of human TAK1a		
CC	(hTKA1a) protein.		
XX	Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;		
SQ			
Alignment Scores:	2.32e-216	Length:	2785
Pred. No.:	3006.00	Matches:	578

[illegible]

```
Db 1123 GACATTGCTTCTCAAAATACGAGTAACAAAGTGACACTAATATGGAGCAAGTTCTCTGCC 1182
Qy 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
Db 1183 ACAATGATACTATTAGCGCTTAGAATCAAAATTTGTGAATAATCAGGCAAGCAACAG 1242
Qy 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerValGluSerLeuPro 380
Db 1243 AGTGAATCTGGAGCTTTAAGCTTGGGAGCCTCCCGTGGGAGCAGTGTGGAGAGCTTGCC 1302
Qy 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
Db 1303 CCAACCTCTGAGGCAAGAGATGAGTGCTGACATGCTCAAAATAGAAGCTAGGATCGCC 1362
Qy 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1363 GCAACCAAGCAACGACGACGACCAAGAGCTAGATCATCCAAAGCTTGACTGTAACGTGA 1422
Qy 421 ThrGluProGlyGlnValSerArgSerSerProSerValArgMetIleThrThr 440
Db 1423 ACAGAACCTGCTGAGTGAGCAGTAGGTCTATCCAGTCCCAGTGTCCAGAATGATTACTACC 1482
Qy 441 SerGlyProThrSerGluLysProThrArgSerHisProThrThrProAspAspSerThr 460
Db 1483 TCAGGACCAACCTCAGAAAAAGCCAACTCGAAGTCAATCCATGGACCCCTGATGATCCACA 1542
Qy 461 AspThrAsnGlySerAspAsnSerIleProMetAlaThrLeuThrLeuAspHisGlnLeu 480
Db 1543 GATACCAATGGATCAGATACTCCATCCCAGTGGCTTATCTTACATGGATCACCACACTA 1602
Qy 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
Db 1603 CAGCCTCTAGCAGCGTCCCAAACTCCAAAGATCTATGGCAGTGTGGAACAGCAATTGT 1662
Qy 501 LysMetIaGlnGluThrMetLysValGlnThrGluIleAlaLeuLeuGlnArgLys 520
Db 1663 AAATGGCAACAAGATATATGAAGTTCAACAGAAATTGCAATGTTATTTACAGAGAAAG 1722
Qy 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
Db 1723 CAAGACTAGTTGAGAACTGGACAGGATGAAGAGCAGCAAAATACATCTCCGCTG 1782
Qy 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyThrGlnGln 560
Db 1783 GTACAGGAACATAAAAGCTTTTAGATGAAACAAAAAGCCTTTCTACTTACTACCAGAA 1842
Qy 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
Db 1843 TGCAAAAAACAACCTAGAGGTCATCAGAAAGTCAGCAGCAGAAACGACAGGCACTTCA 1899

RESULT 7
AAT85094
ID AAT85094 standard; cDNA; 2443 BP.
XX
AC AAT85094;
XX
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1 cDNA.
XX
KW TGF-beta; signal transmission; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase; ss.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 157..1896
FT /tag= a
FT /product= TAK-1
XX
XX JP09163990-A.
XX
```

```
PD 24-JUN-1997.
XX
XX 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX
XX WPI; 1997-380171/35.
XX P-PSDB; AAW27092.
XX
XX DNA encoding transforming growth factor-beta-activated kinase, TAK-1
XX - useful for studying the TGF-beta signal transmission system
XX
XX Claim 2; Page 10-12; 20pp; Japanese.
XX
XX The present sequence encodes mouse transforming growth factor-beta
XX (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
XX TAK-1 protein which is involved in the TGF-beta family signal
XX transmission system. TAK-1, also known as activator of MAPK Kinase
XX (AMK-1), is an enzyme which is activated by TGF-beta and bone
XX morphogenetic protein (BMP) and activates MAPK kinase by
XX phosphorylation.
XX
XX Sequence 2443 BP; 669 A; 567 C; 647 G; 559 T; 1 other;
XX
Alignment Scores:
Pred. No.: 6,25e-215 Length: 2443
Score: 2986.00 Matches: 574
Percent Similarity: 99.14% Conservative: 0
Best Local Similarity: 99.14% Mismatches: 5
Query Match: 99.07% Indels: 0
DB: 18 Gaps: 0
US-09-830-144-2 (1-579) x AAT85094 (1-2443)
Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
Db 157 ATGTCGACAGCGCTCCGCGCCTCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 216
Qy 21 AlaProSerGlnValLeuAsnPheGluIleAspTyrlsGluIleGluValGluGlu 40
Db 217 GCGCGCTCGCAGGCTCTGAACCTTCGAAGAGATCGACTACAGAGAGATCGAGTGAAGAG 276
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysLysLysLysTyrArgAlaLysAspVal 60
Db 277 GTTGTGCGAAGAGGAGCTTTTGGAGTAGTTTGCAGAAAGCTAAGTGGAGAGCAAAAGATGTC 336
Qy 61 AlaIleLysGlnIleGluSerGluSerGluSerGluSerGluSerGluSerGluSerGlu 80
Db 337 GCTATTAAACAGATAGAAAGTGAGTCTGAGAGAGAGGCTTTTCATTGTGGAGCTCCGCGAG 396
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrlsGlyAlaCysLeuAsnProVal 100
Db 397 TTGTCGCGTGTGAACCTCTTAACTTGTCAAGTTGTACGAGGCTCGCTGAATCCAGTA 456
Qy 101 CysLeuValMetGluTyrlsAlaGluGlyGlySerLeuTyrlsAsnValLeuHisGlyAlaGlu 120
Db 457 TGTCTTGTGATGAATATGACAGAGGGGGGCTCATTTATATATATGCTGCTGCTGCTGAA 516
Qy 121 ProLeuProTyrlsThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db 517 CCATTGCGCTTACTACACTGCTGCATGCCATGAGCTGGTGTTTTACAGTGTTCACAGGA 576
Qy 141 ValAlaTyrlsLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 577 GTGGCTTACCTGCACAGCATGCAGCCCAAGCGCTGATTACAGGAGACCTCAAGCTTCCA 636
Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 637 AACTTGCTGCTGGTTGCAGGAGGAGCAGTCTCTAAAAATCTGCGATTTTGGTACAGCTTGT 696
```


QY 181 AspIleGIInThRhiSmetThrAsnAsnLysGlySerAlaIATrpMetAlaProGluVal 200
 Db 697 GACATCCAAACACACATGACCAATAATAAGGAGTGCGCTTGATGGCGCTCGAAGTG 756
 QY 201 PhcGluGlySerAsnThySerGluLysCysAspValPhaSerTrpGlyIleIleuTrp 220
 Db 757 TTTGAAGGTAGCAATTAACAGTGAAGAGTGTATGTCCTCAGCTGGGGATTTATCTCTGG 816
 QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
 Db 817 GAAGTGAATAACAGCGCGGAACCTTCGATGAGATCGGTGGCCACCTTCGAATCAAG 876
 QY 241 TrpAlaValHisAsnGlyThrArgProProleuIleLysAsnLeuProLysProIleGlu 260
 Db 877 TGGGCTGTTCAATATGACATCCAGCACCACTGATCAAAATTATTAAGCCCATTTGAG 936
 QY 261 SerLeuMetThrArgCysTrpSerLysAspProSerGluArgProSerMetGluGluIle 280
 Db 937 AGCTGTATGACAGCGTGTGTGTAAAGACCATCTCAGCGCCCTTCAAATGAGGAAATT 996
 QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
 Db 997 GTGAAATATATGACTACTGATCGCGTACTTCCAGGAGCGGATGAGCATTTACGATAT 1056
 QY 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThSerThrGlySerPheMet 320
 Db 1057 CCTTGATGACTCTGTATGAAAGGAGAGACCACTCAGCCACAGACTCGTTCAAG 1116
 QY 321 AspIleIleSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
 Db 1117 GACATTGCTTCTAACAATACAGAAATAAAGTGAACAATAATGGAACAGGTTCTTGCC 1176
 QY 341 ThrAsnAspThrIleLysArgLeuGlySerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
 Db 1177 ACAAAACGACACTATTAAGCCTTGAGTCAAACTGTGTAAGAAACAGGAAAGCAACAG 1236
 QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
 Db 1237 AGTGAATCTGAGCGCTGAGCTTGAGGAGCTCTCGTGGGAGCAGTGTGAGAGACTTGCC 1296
 QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
 Db 1297 CCCACTTCCAGGAGCAAGAGATGAGTCTGACATGCTGAATAAGCCAGGATCGTG 1356
 QY 401 AlaThrThrGlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGly 420
 Db 1357 GCGACTGCAGTAAACGGGCAACAGCGGTAGATCCATCCAAAGCTTGACTGTATCTGGG 1416
 QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
 Db 1417 ACAGAACTGTGTCAGGTGAGACACCGGTCTCATCCAGCCCTTAGTGTGCAATGATCACTACC 1476
 QY 441 SerGlyProThrSerGluLysProThrArgSerHisProThrProAspAspSerThr 460
 Db 1477 TCGAGGACCAACCTCAGGAAGCCAGCTCGCAGTCACCAATGAGCCCTGTATGATTCACA 1536
 QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
 Db 1537 GACACCAATGGCTCAGTAATCTCATCCCAATGGCGTATTTACATCGATCCAGCGCTA 1596
 QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
 Db 1597 CAGGCTCTAGCGCGTGCCCAAACCTCCAAAGAAATCCATGCGCATGTTTCGAACAGCACTGT 1656
 QY 501 LysMetAlaGlnGlyTrpMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys 520
 Db 1657 AAAATGCAACAGAGATATGAAAGTTCAAAACCAATTCGATTTACTACAGAGAAGAG 1716
 QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
 Db 1717 CAAGAACTAGTTGCAGAATTGACACAGATGAAAAAGACGACAAATATCATCTTCGTTCTG 1776

QY 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
 Db 1777 GTACAGGAACATTAATAAGCTTTATGATGAAAAAGCCCTTCTACTTATTTACAGCAA 1836
 QY 561 CysLysLysGlnLeuGlnValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
 Db 1837 TGCAAAAAACACTAGAGGTATCATCAGAACGCAACGACGAAGACAGACACTTCA 1893
 RESULT 8
 AAX99697
 ID AAX99697 standard; cDNA to mRNA; 2866 BP.
 XX
 AC AAX99697;
 XX
 DT 18-OCT-1999 (first entry)
 XX
 DE Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.
 XX
 KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1b; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 163..1983
 FT FT /*tag= a
 FT FT /product= "hTAK1b"
 XX
 PN MO9940202-A1.
 XX
 PD 12-AUG-1999.
 XX
 PF 02-FEB-1999; 99MO-JP00422.
 XX
 PR 30-OCT-1998; 98JP-0309316.
 PR 06-FEB-1998; 98JP-0026003.
 XX
 PA (TANA) TANABE SEIYAKU CO.
 XX
 PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;
 XX
 DR WPI; 1999-494298/41.
 DR P-PSDB; AAY28997.
 XX
 PT Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases
 XX
 PS Examples; Page 39-43; 49pp; Japanese.
 XX
 CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the nucleotide sequence of human TAK1b
 CC (hTAK1b) protein.
 XX
 SQ Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 1,41e-214 Length: 2866
 Score: 2982.50 Matches: 578
 Percent Similarity: 95.38% Conservative: 0
 Best Local Similarity: 95.38% Mismatches: 1
 Query Match: 98.95% Indels: 27
 DB: 20 Gaps: 1
 US-09-830-144-2 (1-579) x AAX99697 (1-2866)


```
Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20
    |||||
Db 163 ATGCTACAGAGCTCTGCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGGCGGTGAGATCGAA 222
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrIysGluIleGluValGluGlu 40
    |||||
Db 223 GCCCCTTCCCAGGTCTCAACTTTGAAGAGATCGACTACAAGGAGATCGAGGTGGAAGAG 282
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysIysAlaIysTrpArgAlaLysAspVal 60
    |||||
Db 283 GTTGTGGAAGAGAGAGCTTTGGAGTTGTTTGCAAAAGCTAAGTGGAGAGCAAAAAGATGTT 342
Qy 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
    |||||
Db 343 GCTATTAAACAAATAGAAAGTAATCTGAGAGGAAACGGTTTATTGTAGAGCTTCCGCAG 402
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
    |||||
Db 403 TTATCCCGTGTGAACCATCTTAATATTGTAAGGCTTTATGGAGGCTGCTTGAATCCAGTG 462
Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
    |||||
Db 463 TGTCTTGTGATGGAAATATGCTGAAGGGGGCTCTTTATATATAATGTGTGCATGCTGTGAA 522
Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
    |||||
Db 523 CCATTGCCATTATATACGTGTCGCCACGCAATGAGTTGGTGTTTACAGTGTTCCTCCAAAG 582
Qy 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
    |||||
Db 583 GTGGCTTATCTTCACAGCATGCAACCCCAAGGCGTAATTCACAGGACCTGAAACCA 642
Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
    |||||
Db 643 AACTTACTGTGTGTGAGGGGGACAGTCTTAAAAATTTGTGATTTTGGTACAGCCTGT 702
Qy 181 AspileGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
    |||||
Db 703 GACATTGAGACACACATGACCAATAAACAGGGGAGTGCTGCTTGGATGGCACCTGAAGTT 762
Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
    |||||
Db 763 TTTGAAGGTAGTAATACAGTGAANAATGTGACGTCTTCAGCTGGGGTATTATCTTTGG 822
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyProAlaPheArgIleMet 240
    |||||
Db 823 GAAGTGATAACGCGTCGAAAAACCTTTGATGAGATTGGTGGCCAGCTTTCGAATCATG 882
Qy 241 TrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeuProLysProIleGlu 260
    |||||
Db 883 TGGGCTGTTTCAATAATGGTACTCGACCACTCATATAAAAAATTTACCTAAGCCCATTTAG 942
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
    |||||
Db 943 AGCCTGATGACTCGTTGTTGGTCTAAAGATCCCTTCCAGCGCCCTTCAATGGAGGAAAT 1002
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
    |||||
Db 1003 GTGAAAATAATGACTCACTTGATCGGTACTTTCAGGAGACAGATGAGCCATTACAGTAT 1062
Qy 301 ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320
    |||||
Db 1063 CCTTGTGAGTATTGAGATGAAGACAGACGCAACTCTGCCACCATGACAGGCTCATTTATG 1122
Qy 321 AspileAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla 340
    |||||
Db 1123 GACATTGCTTCAAAATACGAGTAACAAAGAGTGACACTAATAATATGAGCAAGTTCCTGCG 1182
Qy 341 ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
    |||||
Db 1183 ACAATGATACTATTAGCGCTTAGAATCAAAATGTTGTAANAATCAGGCAAGCAACAG 1242
Qy 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
```

```
Db 1243 AGTGAATCTGGACGTTTAAAGCTTTGGAGCCTCCCTGGGAGCAGTGTGAGAGCTTGCCC 1302
Qy 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
    |||||
Db 1303 CCAACCTCTGAGGGCAAGAGGATGAGTCTGCATGCTCTGAATAGAAAGCTAGGATCGCC 1362
Qy 401 AlaThrThr----- 403
Db 1363 GCAACACACAGCCTATTCCAAGCCTAAACGGGGCCACCCTGAAAACTGCTTCTTTGGCAAC 1422
Qy 404 -----GlyAsnGlyGlnProArgArgArgSerIle 413
Db 1423 ATTCTGGAGTGTCCCTGAGATCGTCATATCAGGCAACGACAGCCCAAGACGTAGATCCATC 1482
Qy 414 GlnAspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerSerPro 433
Db 1483 CAAGACTTGACTGTAACTGGAAACAGAACTGGTCAGGTGAGCAGTAGGTCTATCCAGTCCC 1542
Qy 434 SerValArgMetIleThrSerGlyProThrSerGluLysProThrArgSerHisPro 453
    |||||
Db 1543 AGTGTGAGAAATGATTACTACCTCAGGACCAACCTCAGAAAAAGCCAACCTCGAAGTCATCCA 1602
Qy 454 TrpThrProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyr 473
    |||||
Db 1603 TGGACCCCTGATGATTCCACAGATACCAATGGATCAGATAACTCCATCCCAATGGCTTAT 1662
Qy 474 LeuThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMet 493
    |||||
Db 1663 CTTACACTGGATCACCACCTACAGCCTCTAGCACCGTCCCAAACTCCAAAGAATCTATG 1722
Qy 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
    |||||
Db 1723 GCAGTGTTTGAACAGCAITTTGAAAAATGSCACAGAATATATGAAAGTTCAACAGAAAT 1782
Qy 514 AlaLeuLeuLeuGlnArgLysGlnGluLeuValAlaGluLeuAspGlnAspGluLysAsp 533
    |||||
Db 1783 GCATTGTTATTACAGAGAAAGCAAGAACTAGTTGCAGAACTGGACCAGGATGAAAGGAC 1842
Qy 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSer 553
    |||||
Db 1843 CAGCAAAATACATCTCGCCTGTGACAGGAACATAAAAAAGCTTTTAGATGAAAAACAAAAGC 1902
Qy 554 LeuSerThrTyrTyrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGln 573
    |||||
Db 1903 CTTTCTACTTACTACCAGCAATGCAAAAAACAACCTAGAGGTCAATCAGAAAGTCAGCAGCAG 1962
Qy 574 LysArgGlnGlyThrSer 579
    |||||
Db 1963 AAACGACAAAGGCACCTTCA 1980
RESULT 9
AAx99698
ID AAx99698 standard; cDNA to mRNA; 1704 BP.
XX
AC AAx99698;
XX
DT 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 1..1704
FT /*tag= a
FT /product= "hTAK1c"
XX
```


QY 494 AlaValPheGluGlnHisCysLysMetAlaGlnGluTyrMetLysValGlnThrGluIle 513
Db 1524 ----- 1524
QY 514 AlaLeuLeuGlnArgLysGlnGluValAlaGluLeuAspGlnAspGluLysAsp 533
Db 1525 -----CAAGAACTAGTTGCAAGAACTGGACAGATGAAAGGAC 1563
QY 534 GlnGlnAsnThrSerArgLeuValGlnGluHisLysLysLeuAspGluAsnLysSer 553
Db 1564 CAGCAAAATACATCTCGCCTGGTACAGGACATAAAAAGCTTTAGATGAAAACAAAGGC 1623
QY 554 LeuSerThrTyrThrGlnGlnCysLysLysGlnLeuGluValIleArgSerGlnGlnGln 573
Db 1624 CTTTCTACTTACTACAGCAATGCAAAAACAACTAGAGTGCATCAGAGTGCAGCAGCAG 1683
QY 574 LysArgGlnGlyThrSer 579
Db 1684 AAACGAAAGGCACCTTCA 1701

RESULT 10

ABL02489

ID ABL02489 standard; cDNA; 3367 BP.

XX ABL02489;

AC ABL02489;

XX 26-MAR-2002 (first entry)

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 1949.

DE Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

KW Drosophila melanogaster.

OS WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-191637P.

XX 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

PA Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

XX P-FSDB; ABB58386.

DR New isolated nucleic acid detection reagent for detecting 1000 or more

DR genes from Drosophila and for elucidating cell signalling and cell-cell

DR interactions -

XX Claim 1; SEQ ID NO 1949; 21bp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA

XX sequences (ABL01840-ABL16175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;

XX Alignment Scores:

Pred. No.: 2,66e-63 Length: 3367
Score: 971.00 Matches: 232
Percent Similarity: 52.22% Conservative: 97
Best Local Similarity: 36.83% Mismatches: 202
Query Match: 32.22% Indels: 100
DB: 23 Gaps: 14

US-09-830-144-2 (1-579) x ABL02489 (1-3367)

QY 17 GluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIle 36
Db 932 GAGATGCCACAGCATCGCTGGACGCACATCGAGCAGCCTATGTGGACTTTCAGTGAGATA 991
QY 37 GluValGluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArg 56
Db 992 ACATAAGAGAGAAAGTCGCCCATGGGTCTCTACGAGTGTGTGCAAGGCCGTTTGGCGC 1051
QY 57 AlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal 76
Db 1052 GACAAGCTGGTTGCCGTCAAGGAGTTCTTCGCCAGCGCCGAGCAGAGGACATCGAGAAG 1111
QY 77 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLysLeuTyrGly---- 94
Db 1112 GAGGTGAAGCAGTGTCTCGCGCTGAAGCACCAGCAATCATCTCTGCACGGGATATCC 1171
QY 95 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 114
Db 1172 TCGTACCAGCAGCGCCACCTACCTGATTAATGGAGTTGCCGAAGGTGGATCGCTGCAAC 1231
QY 115 ValLeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTyrCys 134
Db 1232 TTCCTTCAACGCG--AAGGTGAAGCGCGCATATTCCTGGCCCCACGCCATGAGCTGGCG 1288
QY 135 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 154
Db 1289 CGCCAATGTGCAGAGGGTCTGGCATATTTGCATGCCATGACGCCAAACACCATTAATACAT 1348
QY 155 ArgAspLeuLysProProAsnLeuValLeuValAlaGlyGlyThrValLeuLysIleCys 174
Db 1349 CGCGACGTGAAGCGCTGAACCTGCTTTCACCAACAGGAGCGCAATCTGAAGATATGC 1408
QY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1409 GACTTCGGCACGGTGGCGACAAGTCGACCATGATGACCAACAATCGCGCAGTGGCGCT 1468
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 1469 TGGATGGCGCCCGAGGTCTCGAAGGCTCCAAGTATACGAGAAAGTGTGACATTTTATAGC 1528
QY 215 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 1529 TGGGCCATTGTTCTATGGGAGGTTCGTCCAGGAACAGCCCTTTAAAGGCATCGCAAT 1588
QY 235 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1589 ---GCCTACACCATCCAGTGGAAAGTCTACAAGGGTGAACGCCCGCGCTGTGACCACT 1645
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArg 274
Db 1646 TGCCCCAAGCGCATCGAGACCTGATGACCGCTGCTGGAAACCGTGCCTCGAGGATCGC 1705
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 294
Db 1706 CCGTCATCAGTACATAGTGGCGTTATGCAGAGATCGTCAAGGACTATATACGGGGGGG 1765
QY 295 AspGluProLeuGlnTyr----- 300
Db 1766 GACAAGGCCCTGGAATACACGTTTGTATCAACAGATTGTCCACCAAGAGAGACGCGC 1825
QY 301 -----ProCysGlnTyrSerAspGluGlyGlnSerAsn----- 311
Db 1826 ACGGTGGCGCGCTCAACCGGATAGCCTCAGTTTCGAGGAGGGGGAACATGAGCCCTCGTCC 1885

OY 312 -----SerAla 313
 Db 1886 ACACAGTTAACCGACACCGCGCCCAACGCCAATGTGAACGCGATGATCAATAA 1945
 OY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 Db 1946 ACACAGCTAGCTCAATGACCGAATAATACCTATCATCAATCATCGACATCGACCGACG 2005
 OY 334 AsnMetGluGlnValPro-----AlaThrAsn-----AspThr 344
 Db 2006 AACTCGGGCCACTGTGACATATATCCGCTATTCTTATATGTCACCAATCGCTGGACCGC 2065
 OY 345 IleLysArgLeuGluSerLysLeuLysAsn----- 355
 Db 2066 ATTCGCCAGAGAGAGAGACGACGAGCCGAAAGCATAGCTTCACCTCACCTCGTGCGCT 2125
 OY 356 GlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSer 375
 Db 2126 GAGGCCACTACCGCTCGAATCCGGAACGCGCATGATCTGATGCGCTGCAAGCCC 2185
 OY 376 ValGluSerLeuProProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIle 395
 Db 2186 ATGGAGCACTC-----ACCTTCACGTGAGCGGAATGCTTTGATGTAGTCCACG 2239
 OY 396 GluAlaArgIleAlaAlaThrThrGlyAsnGlyLysProArgArgSerIleGlnAsp 415
 Db 2240 GAAAGCAGCAGCAGCAGCAGCAACGAAAGCGATGCGCCGCAACGA----- 2287
 OY 416 LeuThrValThrGlyThrGluPro----- 423
 Db 2288 CTCACCGTACGACACGACCAAGCGGTGATGATGACCAACGATCTGCCAACACAGCGC 2347
 OY 424 Gly-GlnValSerSerSerSerSerSerProSerValArgMetIleThrThrSerGlyPr 443
 Db 2348 GGCATCCAGCCCACTCGAAGCAGACGCGATGCGCATGCAATGCAATGTT-----GGCA 2398
 OY 443 oThrSerGluLysProThrArgSerHisPro-----TrpThr 455
 Db 2399 GCMAAGATGAGAGAGCTGACGAGGCAAGAGCATGACAGAGATTTGTCACTCGTTGGAC 2458
 OY 455 rProAspAspSerThrAspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThr 475
 Db 2459 GTC-GAGCTGGATCCGACGAGATGAGACGACGACCGAACGACTCAGCCGAGAT 2517
 OY 475 rLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMetAlaVal 495
 Db 2518 TCTTGATCCGAGACTCGACGACGAGCCGCGATATCCCAAGATGCGCATCCAGCTCAT 2577
 OY 495 lPheGluGlnHisCysLysMetAlaGlnGlyTyrMetLysValGlnThrGluIleAlaLe 515
 Db 2578 CTACCGGAGACCGACGACATGCGCAAGAGTACTGACGCTGACACGACCAACCTTACTA 2637
 OY 515 uLeuLeuGlnArgLysGlnGlnLeuValAlaGluLeuAspGlnAspGluLysAspGlnI 535
 Db 2638 CGGCGACGAGACTTAAAGACAGACTCATGTCGATGATGACCGACCGACCGACGACAGA 2697
 OY 535 nAsnThrSerArgLeuValGlnGlnHisLysLysLeuLeuAspGluAsnLysSerLeuSe 555
 Db 2698 G-----CAGGAGCTTCTGCGCAAGATGAAGACAAAGAGGGTCTTCA 2739
 OY 555 rThrTyrTyrGlnGlnCysLysLysGln 564
 Db 2740 GAGTCTTTACACATCTGCAGCAGAG 2767
 RESULT 11
 ID ABL02488 standard, cDNA, 10997 BP.
 AC ABL02488;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.

XX
 KW Drosophila, developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 PN MO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PMD, Myers EW;
 DR WPI, 2001-656860/75.
 DR P-PSDB; ABB58385.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 1946; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB116175), expressed DNA
 CC sequences (AB101840-AB116175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;
 Alignment Scores:
 Pred. No.: 1 84e-42 Length: 10997
 Score: 702.50 Matches: 190
 Percent Similarity: 41.09% Conservative: 73
 Best Local Similarity: 29.69% Mismatches: 164
 Query Match: 23.31% Indels: 214
 DB: Gaps: 14
 US-09-830-144-2 (1-579) x ABL02488 (1-10997)
 OY 25 ValLeuAsnPhedGluGluIleAspTyrLysGluIleGluVal----- 38
 Db 2469 GTAATTGATATTGAGAGATTTCACACAGAGGATACACTGTTGATCAGATCTAATA 2528
 OY 39 -----GluGluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrp 55
 Db 2529 TCTTTTCCCGCAGAAAGTCGCGCATGGGCTCTACGGAATGGTGTGCAAGGCCGTTGG 2588
 OY 56 ArgAlaLysAspValAlaIleLysGlnIleGluSerLysSerGluArgLysAlaPheIle 75
 Db 2589 CGCGACAGAGCTGGTGGCTGTAAGAGTTCTTCCACGCCCGACGACGACATGACG 2648
 OY 76 ValGluLeuArgGlnLeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGly--- 94
 Db 2649 AAGAGGTGAAGACAGTTGTGCGCGTGAAGACCCGAACATCATCGCTGACGCGGATA 2708
 OY 95 ---AlaCysLeuAsnProValCysLeuValMetGluTyrTrpAlaGluGlyGlySerLeuTyr 113
 Db 2709 TCCTGTACACAGACGCGCACCTACTGATATGAGATGTCCGCGAAGGTGATCGCTCAC 2768
 OY 114 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 133

XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX
PS Claim 1; SEQ ID 8355; 44pp; English.

XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70110), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.

XX
SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;

Alignment Scores:
Pred. No.: 2,13e-27 Length: 3454
Score: 493.50 Matches: 145
Percent Similarity: 46.49% Conservative: 87
Best Local Similarity: 29.06% Mismatches: 191
Query Match: 16.37% Indels: 79
DB: 24 Gaps: 14

US-09-830-144-2 (1-579) x ABL70018 (1-3454)

QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
DB 535 GCACCCGGCGGCTCCAGCTGCGCCAGGAGATCCCTTCCAGAGCTGCAGCTAGGAGG 594
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
DB 595 ATCATCGGTGTGGGGGCTTTGGCAAGGTCTATCGGGCCCTGTGGCGTGGCGAGGAGTG 654
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
DB 655 GCAGTCAAGGCGCGCCGGCTGGACCTGTGAGAGGACCCGGCAGTGACAGCGGAGGAGTG 714
QY 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
DB 715 TGCCAGGAGCCCGGCTCTTTGGAGCCCTGCAGACCCCAACATAATTGCCCTTAGGGGC 774
QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
DB 775 GCCTGCTCAACCCCGCCACACCTCTGCCTAGTGTGATGATGCCGGGGTGTGTCACCTG 834
QY 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
DB 835 AGCAGGGTGTGTCAGGTGCGCGGGTGCCACCTCACGTG-----CTGGTCAAC 882
QY 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
DB 883 TGGGCTGTGAGGTGGCGGGGCATGAACCTACTACATGATGCCCTGTGCCCATC 942
QY 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
DB 943 ATCCACGGGACCTCAAGTCCATCAACATCTGATCTGTCGAGGCCATCAGAACACACAC 1002
QY 166 AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr 184
DB 1003 CTGCGACACACGGTGTCAAGATCAGGACTTCGGCTCGCCCGGAGTGCGCACAGACC 1062

QY 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
DB 1063 ACAACATGAGCGCTCGGGGACCTACGCTGGATGGCGCGAGGTTATCGCTCTCTCC 1122
QY 205 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
DB 1123 CTCTTCTCCAAAAGCAGTGTCTGGAGCTTCGGGGTCTGCTGTGGAGCTGTGACG 1182
QY 225 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal--- 243
DB 1183 GGGAGGTCCTTACCTACCGTGAGATC-----GACGCTTGGCGTGGCGTATGGCGTGGCT 1236
QY 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263
DB 1237 ATGAATAAGCTACGCTGCCATTCCTCCAGTCCCGCGAGCCCTTGGCCGCTCTCTG 1296
QY 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValIle 283
DB 1297 GAGGAATGCTGGGACCCAGACCCCGCGGCGCCAGATTTCGGTAGCATCTTGAAGCGG 1356
QY 284 MetThrHisLeuMetArgTyrPhePro----- 292
DB 1357 CTTGAAGTCAATGACAGACATC-AGCCCTGTTCCAGATGCCATGGAGTCTTCCACTCGCT 1415
QY 292 ----- 292
DB 1416 GCAGGAAGCTGGAAGCTGGAGATTCAGCACATGTTTGTGATGACCTTCGGACCAAGAGAA 1475
QY 293 GlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSer 312
DB 1476 GGAGCTTCGAGCGGTG-AGGAGGAGCTGTCGGGGCGGCACAGGAGCAGCGCTCCAGG 1534
QY 313 AlaThrSerThrGly-----SerPheMetAspIleAlaSerThrAsnThrSer 328
DB 1535 AGGAGCAGCTGCGCGCGGCGGAGCAGGAGTGGCAGAACGCTGAGATGGACATCTGGAAC 1594
QY 329 AsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysAspLeu 348
DB 1595 GGGAGCTGCACCTGCTC-ATGTCCAGCTGACGACGAGAGAGCCCGGGTCCCAAGCGC 1653
QY 349 GluSerLysLeuLysAsnGlnAlaLysGlnGlnSerGluSerGly-----Arg 365
DB 1654 AAGGGCAACTTCAAGCGCAGCCGCTGCTCAAGCTGGGGAAGCGGCGGACACATCAGC 1713
QY 366 LeuSerLeuGlyAlaSerHisGlySerValGluSerLeuProProThrSerGluGly 385
DB 1714 CTGCCCTCTGCTTTGAGCATAAAGATCACAGTCCAGGCTCTCCAACTCTGTGAT----- 1767
QY 386 LysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAlaThrThrGlyAsn 405
DB 1768 AAGCGGAAGGATCCGATGGGCGCCAGCCCTTCGCAAGCCCGCAGCATCATC----- 1818
QY 406 GlyGlnProArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGln 425
DB 1819 -----CCCGCGCTGAGGCGCATTCGC-----CTGACTCCCGTGGACTGTGGTGC 1863
QY 426 ValSerSerArgSerSerProSerValArgMetIleThrThrSerGlyProThrSer 445
DB 1864 AGCAGCAGTGCACGACGAGTGGAGGAAGTGGACATGGAGCGCGGTGGGCGCCCAAG 1923
QY 446 GluLys-----ProThrArgSerHisProThrThrProAspAspSer 459
DB 1924 AAGGAAGAACTGTCGGGGGCAAGAGAGGAGCAACGTCGGGGGCCAGCTCCACC 1980
RESULT 14
AAZ99726
ID AAZ99726 standard; cDNA; 2120 BP.
XX
XX AAZ99726;
AC
XX
XX 12-JUL-2000 (first entry)
DT
XX cDNA encoding human cardiovascular system associated protein kinase-2.
DE


```
||||: ||| ||| |||: ||| |||: ||| |||: |||
Db 1133 ACAAGTACGGGAGGCGCATGGCATGAACCAAGTCTCGAGCCATGATGCTGATGGC 1192
Qy 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGATATCTTCTCAATGAACAAGAGGAGCTGTG 1231
RESULT 15
AAA75674
ID AAA75674 standard; DNA; 1706 BP.
XX
AC AAA75674;
XX
DT 22-JAN-2001 (first entry)
XX
DE DNA encoding a human regulator of intracellular phosphorylation.
XX
KW Human; intracellular phosphorylation regulator; HRIP; stroke; myeloma;
KW neurological disorder; Parkinson's disease; demyelinating disease;
KW meningitis; developmental disorder; neuromuscular disorder; cancer;
KW myasthenia gravis; cell proliferative disorder; actinic keratosis;
KW arteriosclerosis; atherosclerosis; leukaemia; melanoma; bronchitis;
KW autoimmune disorder; inflammatory disorder; Addison's disease;
KW acquired immunodeficiency disease; allergy; diabetes mellitus;
KW rheumatoid arthritis; microbial infection; trauma; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 138..1505
FT CDS /tag= a
FT /product= "regulator of intracellular phosphorylation"
XX
XX WO200055332-A2.
XX
XX 21-SEP-2000.
XX
XX 17-MAR-2000; 2000WO-US07277.
XX
XX 18-MAR-1999; 99US-0125593.
XX
XX 20-MAY-1999; 99US-0135049.
XX
XX 09-JUL-1999; 99US-0143188.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandnan O, Tang YT, Yue H, Hillman JL, Baughn MR, Azimzai Y;
XX Lu DAM, Au-Young J;
XX
XX WPI; 2000-602121/57.
XX
XX P-PSDB; AAB18657.
XX
XX Novel human intracellular phosphorylation regulator polypeptides and
XX polynucleotides for diagnosis, prevention and treatment of
XX neurological, cell proliferative and autoimmune/inflammatory disorders
XX
XX Claim 4; Page 88; 96pp; English.
XX
XX The present sequence encodes a human regulator of intracellular
XX phosphorylation (HRIP). HRIP is useful for screening agonists and
XX antagonists of HRIP polypeptide. HRIP and its agonist or antagonist
XX are useful for treating a disease or condition associated with
XX decreased or increased expression of functional HRIP. Diseases treated
XX or diagnosed include neurological disorders such as stroke, Parkinson's
XX disease, demyelinating diseases, bacterial and viral meningitis and
XX other developmental disorders of the central nervous system,
XX neuromuscular disorders, myasthenia gravis, cell proliferative disorders
XX such as actinic keratosis, arteriosclerosis, atherosclerosis and cancer
XX including leukaemia, melanoma, myeloma and cancer of the adrenal gland,
XX bladder, bone, bone marrow, liver, lung, muscle, ovary, autoimmune/
XX inflammatory disorder such as Addison's disease, acquired
XX immunodeficiency diseases, allergies, bronchitis, diabetes mellitus,
XX rheumatoid arthritis, microbial infection and trauma.
```

```
XX
SQ Sequence 1706 BP; 509 A; 349 C; 427 G; 421 T; 0 other;
Alignment Scores:
Pred. No.: 6.3e-27 Length: 1706
Score: 482.00 Matches: 134
Percent Similarity: 48.75% Conservative: 81
Best Local Similarity: 30.39% Mismatches: 156
Query Match: 15.99% Indels: 70
DB: 21 Gaps: 18
US-09-830-144-2 (1-579) x AAA75674 (1-1706)
Qy 20 GluAlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluLeuGluValGlu 39
Db 135 GAGATGCGTCTCTCGGTGCTCTTTGTGCAAAATAAATTTGATGACTTGCGAGTTT 194
Qy 40 GluValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrp-----ArgAla 57
Db 195 GAAAACTCGCGTGGAGGAAGTTTTGGGAGTGTTCGAGCCAAATGGATATCACAGGAC 254
Qy 58 LysAspValAlaIleLys-----GlnIleGluSerGluSerGluArgLysAlaPhe 74
Db 255 AAGAGGTGCTGTAAAGAAAGCTCTCAAAATAGAGAAAGAGGCAGAA----- 302
Qy 75 IleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
Db 303 -----ATACTCAGTGTCTCTCAGTCACAGAAACATCATCCAGTTTATGGA 347
Qy 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
Db 348 GTAATTTCTGAACCTCCCAACTATGCGCAATGTGCAGAAATATGCTTCTCGGATCATC 407
Qy 113 TyrAsnValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAla 130
Db 408 TATGATTACATTAAACAGTAACAGAGTGAGGAGATG-----GATATGGATCATT 458
Qy 131 MetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLys 150
Db 459 ATGACCTGGGCCACTGATGTAGCCAAAGGAATGCATTATTACATATGGAGGCTCTCTGC 518
Qy 151 AlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
Db 519 AAGGTGATTACAGAGACCTCAAGTCAAGAAACGTTGTATAGCTGTGTATGGA---GTA 575
Qy 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys 190
Db 576 TTGAAGATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACACACATGTCC 632
Qy 191 -----GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu 208
Db 633 TTGGTTGGAACCTTTCCCATGGATGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTGAGAA 692
Qy 209 LysCysAspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysPro 228
Db 693 ACTGTGACACATATCTCTGTGTGTGTCTCTGGGAGATGCTAAACAGGAGGAGTCCCC 752
Qy 229 PheAspGluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThr 247
Db 753 TTTAAAGGTTTGGAAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACGAG 806
Qy 248 ArgProProGluIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrp 267
Db 807 AGATTAAACATTCCAAGCAGCTTGCCCAAGAGTTTGTGTAACCTGTTCATCATCAGTGTGG 866
Qy 268 SerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeu 287
Db 867 GAAGCTGATGCCAAGAACCGCCATCATCAAGCAAAATCATTTCAATCACTCTG----- 917
Qy 288 MetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGlu 307
Db 918 -----GAG 920
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 22:59:44 ; Search time 1808.11 Seconds
(without alignments)
5186.195 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKQLEIVRSQQKROQTS 579

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09830144/runat_04122002.141353.2264/app.query.fasta.1.1422
-DB=EST -QFMT=faststep -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @cgn.1.1.2441 @runat_04122002.141353.2264 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: em_estba:*
2: em_esthum:*
3: em_estcin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_oth:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1443.5	47.9	956	14	BQ958174
2	1379	45.8	1062	13	BM554120
3	1242.5	41.2	998	9	AL525728
4	1225	40.6	771	13	BI093821
5	1212	40.2	739	13	BI696710
6	1140	37.8	768	13	BI868773
7	1119	37.1	827	9	AL550589
8	1116	37.0	688	13	BJ062988
9	1116	37.0	696	13	BJ074867
10	1116	37.0	1006	14	BQ219348
11	1112	36.9	663	10	BB617944
12	1086	36.0	646	13	BJ073883
13	1079	35.8	604	13	BM426610
14	1063	35.3	929	9	AL520975
15	1061	35.2	686	12	BG548917
16	1055.5	35.0	868	12	BE742348
17	1047	34.7	858	13	BG916821
18	1009.5	33.5	910	12	BF780358
19	1006	33.4	641	14	BQ386875
20	995	33.0	621	10	BB618631
21	994.5	33.0	1054	14	BQ067693
22	992	32.9	711	10	AV718168
23	955	31.7	968	12	BG027707
24	954	31.7	599	10	BE284386
25	943.5	31.3	681	12	BF678411
26	923.5	30.6	941	12	BG029948
27	923	30.6	551	9	AA461278
28	921	30.6	747	13	BI334774
29	920	30.5	542	10	AW960377
30	912	30.3	822	13	BI757369
31	901.5	29.9	586	13	BJ035129
32	880.5	29.2	760	12	BG773488
33	873	29.0	682	10	AV398933
34	866	28.7	602	10	BE080188
35	852.5	28.3	745	13	BI767913
36	843	28.0	681	10	AV398935
37	827.5	27.5	1191	13	BM472107
38	814	27.0	1033	12	BE882182
39	811.5	26.9	673	9	AJ392980
40	809	26.8	647	10	BB665645
41	794.5	26.4	594	13	BJ014293
42	782	25.9	729	12	BG826184
43	750	24.9	671	12	BG700279
44	741.5	24.6	585	9	AL773651
45	740.5	24.6	718	9	AL652289

ALIGNMENTS

RESULT 1
BQ958174
LOCUS
DEFINITION AGENCOURT_10014737 NIH_MGC_101 Homo sapiens CDNA clone
IMAGE:6497311 5', mRNA sequence.
ACCESSION BQ958174
VERSION BQ958174.1 GI:22373652
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 956)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLM2678 row: c column: 08
High quality sequence stop: 686.
Location/Qualifiers

FEATURES

source

1..956
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:6497311"
/clone_lib="NIH_MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site 1: EcORI; Site 2: XhoI. cDNA made by oligo-dT priming. Directionally cloned into EcORI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by Ling Hong in the Laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 330 a 222 c 199 g 204 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 5,97e-126 Length: 956
Score: 1443.50 Matches: 290
Percent Similarity: 95.10% Conservative: 1
Best Local Similarity: 94.77% Mismatches: 11
Query Match: 47.89% Indels: 4
Gaps: 2

US-09-830-144-2 (1-579) X BQ958174 (1-956)

QY 263 MetThArgCysTrpSerIysAspProSerGlnArgProSerMetGlnGluIleValIys 282
DB 2 ATGACTCGTGTGGTCTTAAGATCCTCCAGCGCCCTTCATGAGAAATGTGAAA 61
QY 283 IleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnIlyrProCys 302
DB 62 ATTAATGACTGACTTGATCGCGTACTTCCAGGAGCAGATGAGCAATTACGATCTCTGT 121
QY 303 GlnIlyrSerAspGluGlnSerAsnSerAlaThrSerThiGlySerPheMetAspIle 322
DB 122 CAGATTCAGATGAGAGACAGACCAACTGCGCACAGATACAGGCTCATTCATGACATT 181
QY 323 AlaSerThrAsnThrSerAsnIlyrSerAspThrAsnMetGlnGlnValProAlaThrAsn 342
DB 182 GCTTCTCAAAATCGAGTAAACAAAGGACACTTAATATGAGCAAGTCTCCGCCAATA 241
QY 343 AspThrIleIysArgLeuGlnSerIlyrLeuLeuIlyrAsnGlnAlaIysGlnGlnSerGlu 362
DB 242 GATACCTTTAAGCGCTTAGAATCAAATTTGTAAAAATCAGCAACAGACAGATGAA 301
QY 363 SerGlyIlyrGlnSerLeuGlyAlaSerHisGlySerSerValGlnSerLeuProPheThr 382
DB 302 TCTGAGAGTTTAAGCTTGGAGCGCTCCCGTGGAGGACAGTGTGAGAGCTTGGCCCCAAC 361
QY 383 SerGlnGlyIlyrAsnMetSerAlaAspMetSerGluIleGlnAlaArgIleAlaIleThr 402
DB 362 TCTGAGGCAAGAGATGAGATGCTGACATGCTGAATAATAGACTAGATCGCCGCAACC 421
QY 403 ThrGlyAsnGlnIlyrProArgArgSerIleGlnAspLeuThrValThrGlyThrGln 422
DB 422 ACGAGCAACGACAGCAACGATGATTCATCAAGACTTGAAGCTGTAAGTGAACAGAA 481

QY 423 ProGlnIlyrValSerSerArgSerSerProSerValArgMetIleThrThrSerGly 442
DB 482 CTTGGTACGATGAGCAGTAGTATCATCCAGTCCAGTCAAGATGATTAATCTCCAGGA 541
QY 443 ProThrSerGlnIlyrProTrpArgSerHisProTrpThrProAspAspSerThrAspThr 462
DB 542 CCAACCTCCAAAAGCCAACTCCAAAGTCAATGATCCATGAGCCCTGATATTCACAGATAC 601
QY 463 AsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeuGlnPro 482
DB 602 AATGATCAGATTAATCTCATCCCAATGGCTTATCTTACACTGGATCACAACATACAGCT 661
QY 483 LeuAlaProCysProAsnSerIlyrGlnSerMetAlaValPheGlnIlyrHisCysIlyrMet 502
DB 662 CTACACCGGCGCCCAAACTCCAAAGATCTATGACAGTGTGAAACGACTGTGAAAATG 721
QY 503 AlaGlnIlyrMetIlyrValGlnThrGluIleAlaLeuLeuLeuGlnArgIlyrGlnIly 522
DB 722 GCACAAATATATATGAAAGTTCACAAAGCAAGAAATGCATTTGTTATTCAGAGAAAGCA 781
QY 523 LeuValAlaGlnLeuAspGlnAspGlnIlyrAspGlnGlnAsnThrSerArgLeuVal-GI 542
DB 782 CTAGTTCAGAACTGACGACAGATGAAAGGACAGCAACAAATACATCTCGCTCGTAAAC 841
QY 542 ngIuHisIlyrIlyrLeuLeu---AspGluAsnIlyrSerLeu-----SerThrTyrTrpI 559
DB 842 GGAACATTAAGAAAGCTTTTAAATGAAAGCAAAAGCTTTTCTACCTTACCTACAGNC 901
QY 559 ngIuCysIlyrIlyrGln 564
DB 902 AATGCCAAAAAAACAA 917

RESULT 2
BM554120 1062 bp mRNA linear EST 20-FEB-2002
LOCUS
DEFINITION
AGENCOURT 6546903 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742353
5', mRNA sequence.
ACCESSION
BM554120
VERSION
BM554120.1 GI:18793446
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1062)
NIH-MGC http://mgc.nci.nih.gov/.
TITLE
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLM2760 row: f column: 18
High quality sequence stop: 730.
Location/Qualifiers

FEATURES

source

1..1062
/organism="Homo sapiens"
/db xref="taxon:9606"
/clone="IMAGE:5742353"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH10B"
/note="Organ: brain; Vector: PCWV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source: normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for

Db 429 ATTATCCCGGTACACCATCTTAATGTGAAGCTTTATGAGCCTGCTGAATCCACT 488
Qy 100 |CysLeuValMetGluTyr-AlaGluGlySerLeuTyrValLeuHisGlyAlaG 120
Db 489 GTGCTTGATGGAATATAGTGAAGGGGGCTTTAATAATATGCTGACGTGCTG 548
Qy 120 |LurLeuLeuProTyrTyrThAlaAlaHisAlaMetSerTPrpCysLeuGlnCysSerGlnG 140
Db 549 AACCATTTGCATATTAATACGCTGCGCCAGCAATGARTTGTTTACAGTGTTCCTCAAG 608
Qy 140 |ValAlaTyrLeuHisSerMetGlnProValAlaLeuLeuHisArgAspLeuLysProp 160
Db 609 GAGTGGCTTATCTTCACAGATCCAAACCGCTAATTCACAGGACCTGAAACAC 668
Qy 160 roAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaC 180
Db 669 CAACACTACTGCTGTGTGACGGGGGACAGTCTTAAAAATTTGTGATTTGGTACAGCCT 728
Qy 180 yAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaATPMeAlaProGluV 200
Db 729 GTACATTCAGACACATGACCAATTAACAAGGAGTGTCTGCTTGATGGACCTGAAG 788
Qy 200 alPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTPrpGlyIleLeuT 220
Db 789 TTTTGAAGGTAGTATTAAGTGAAGAAAATGTACCTCTTCAGCTGGGTATTAATCTTT 848
Qy 220 rpgIuValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleM 240
Db 849 GGGAGAGTATMG-CGTGGAAACCTTT-GATGAGATGTGTGCCAGCTTCGAAATCA 906
Qy 240 eCTrPalVal-HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
Db 907 TGTGGGCTGTCTTCAATGATGACTCGACCACTGATTAATAATTACCTAAGCCATT- 965
Qy 260 GluSerLeuMetThrArgCysTPrpSerLys 269
Db 966 GAGAGCCTGATGACTCTGTGTGTTAAA 995

RESULT 4
LOCUS B1093821 771 bp mRNA linear EST 20-JUN-2001
DEFINITION 602860601F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5001504 5',
mRNA sequence.
ACCESSION B1093821
VERSION B1093821.1 GI:14512151
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 771)
NIH-MGC http://mgs.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabsb-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM1035 row: b column: 01
High quality sequence stop: 768.

FEATURES
source
Location/Qualifiers
1..771
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5001504"
/clone_1id="NIH_MGC_10"
/cell_line="MG36"
/lab_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5 kb. Library prepared by Life
Technologies."
BASE COUNT 230 a 163 c 187 g 191 t
ORIGIN
Alignment Scores:
Pred. No.: 1,96e-105 Length: 771
Score: 1225.00 Matches: 232
Percent Similarity: 97.92% Conservative: 3
Best Local Similarity: 96.67% Mismatches: 5
Query Match: 40.64% Indels: 0
DB: 13 Gaps: 0
US-09-830-144-2 (1-579) x B1093821 (1-771)
Qy 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys 190
Db 1 CTTAAAAATTTGTGATTTTGTGACAGCTGTGACCTTGACACTTGACACACATGACCAATTAACAAG 60
Qy 191 GlySerAlaAlaATPMeAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 210
Db 61 GGGAGTGTCTGCTTGATGAGGACCTGGAAGTTTGAAGGTATTAACAGTAAATAATGT 120
Qy 211 AspValPheSerTPrpGlyIleLeuTPrpGluValIleThrArgArgLysProPheAsp 230
Db 121 GACGCTTCACCTGGGGGTATTAATCTTGGGAAGTGTATTAACGGCTCGGAACCCCTTGAT 180
Qy 231 GluIleGlyGlyProAlaPheArgIleMetTrpAlaHisAsnGlyTPrpArgPro 250
Db 181 GAGATGTGTGCCAGCTTCGCAATCATGTGGCTGTTCATATGATGACTCGACACCA 240
Qy 251 LeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTPrpSerLysAsp 270
Db 241 CTGATTAATAAATTTAACTTAAGCCATTGAGAGCTGATGACTGCTGTGTCTTAAGAT 300
Qy 271 ProSerGlnArgProSerMetGluGluIleValLysIleMetThrHisLysMetArgTyr 290
Db 301 CCTTCCCGACCGCCCTTCATATGAGGAATGTGAAAATTAATGACTCACTGATGCGGTAC 360
Qy 291 PheProGluValAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSer 310
Db 361 TTCCAGAGCAGATGAGCCATTACAGATCTCTTGACGATTCAGTGAAGGACGAGC 420
Qy 311 AsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLys 330
Db 421 AACTCTGCCACCATGACAGGCTCATTCATGACATTCTCTCAAAATTAAGATTAACAA 480
Qy 331 SerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSer 350
Db 481 AGTGACACTAATATATGAGCAAGTCTCGCACCAATGATATCTAATTAAGCCCTTAAGATCA 540
Qy 351 LysLeuLeuLysAsnGlnAlaLysGlnGlnSerGlyArgLeuSerLeuGlyAla 370
Db 541 AAATCGTTGAAAATACAGGCAAGCAACAGAGTAATCTGACCTTTAACTGGGAGCC 600
Qy 371 SerHisGlySerSerValGluSerLeuProProThrSerGluGlyLysArgMetSerAla 390
Db 601 TCCCTGGGAGCAGTGTGAGAGCTTGCCCCCAAGCTCTGAGGCAAGAGATGAGTCT 660
Qy 391 AspMetSerGluIleGluValArgIleAlaAlaThrThrGlyAsnGlyGlnProArgArg 410
Db 661 GACATGTCTGAATAATGAAGTAGATGCGCGAACACACAGCCTATTCCAAAGCCTTAACGG 720

RESULT 5
B1696710 739 bp mRNA linear EST 18-SEP-2001
LOCUS B1696710
DEFINITION 603347214F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374971 5',
mRNA sequence.
ACCESSION B1696710
VERSION B1696710.1 GI:15659339
KEYWORDS EST.

SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 739)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM11954 row: k column: 04
 High quality sequence stop: 717.
 FEATURES
 Location/Qualifiers
 1..739
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5374971"
 /clone_lib="NCI CGAP Mam2"
 /tissue_type="tumor, biopsy sample"
 /dev_stages="5 months"
 /lab_host="DH10B"
 /note="organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 169 c 188 g 187 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 1e-104 Length: 739
 Score: 1212.00 Matches: 235
 Percent Similarity: 97.93% Conservative: 2
 Best Local Similarity: 97.11% Mismatches: 5
 Query Match: 40.21% Indels: 0
 DB: 13 Gaps: 0
 US-09-830-144-2 (1-579) x BI696710 (1-739)
 QY 83 ArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCysLeuAsnProValCysLeu 102
 Db 3 CGTGTGAACCATCTTAACATTGTCAAGTTGTACGGAGCCTGCTGAATCCAGTATGTCTT 62
 QY 103 ValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeu 122
 Db 63 GTGATGGAATATGCAGAGGGGGCTCATTTGTAATGTCTGATGGTCTGAACCATTTG 122
 QY 123 ProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly-ValAl 142
 Db 123 CCTTACTACACTGCTGCTATGCATGCATGAGTGTGTTCAGATGTTCCCAAGGTAGTGGC 182
 QY 142 aTyrLeuHisSerMetGlnProIlysalaleuIleHisArgAspLeuLysProProAsnLe 162
 Db 183 TTACCTGCACAGCATGCAGCCCAAGCGCTGATTCAAGGAGCCTCAAGCTCCAAACTT 242
 QY 162 uLeuLeuValAlaGlyGlyThrValIleuIlysalaleuIleCysAspPheGlyThrAlaCysaspI 182
 Db 243 GCTGCTGGTTGCAGAGGGGACAGTTCTAAATAATCTCGATTTTGGTACAGCTTGTGACAT 302
 QY 182 eGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheG 202
 Db 303 CCAACACACATGACCCATTAATAAGGAGTGTGCTTGCATGGCCCTCAAGTGTTCGA 362
 QY 202 uGlySerAsnTyr-serGluIysCysAspValPheSerTyrGlyIleIleLeuTyrPGLuVa 222
 Db 363 AGGTAGCAATTACAGTGAAAGAGTGTGATGCTTCAGCTGGGGTATTATCTCTCTGGGAAGT 422

QY 222 IileThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAl 242
 Db 423 GATAACACCCCGAAACCTTCATGATGATCGTGGCCAGCTTTTCAGAAATCATGTGGC 482
 QY 242 aValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLe 262
 Db 483 TGTTTATAATGGCACTCGACCACTGATCAAAAATTTACCTAAGCCCATTCAGAGCTT 542
 QY 262 uMetThrArgCys-TyrSerLysAspProSerGlnArgProSerMetGluGluIleValI 282
 Db 543 GATGACACCTGTGTGTCTAAGGACCCATCTCAGGC-CTTCAATGGAGGAAATTTGTA 601
 QY 282 ySileMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProC 302
 Db 602 AAATATGACTCACTTGTGCGGTACTTCCAGAGCGGATGAGCCATTCAGTATCCTT 661
 QY 302 ySgln-TyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMetAsp 321
 Db 662 GTCAGGTACTCTGATGAAGGGCAGAGCAATCC-AGCACCAGCAGGCTCGTTTCATGAA 720

RESULT 6
 BI688773 768 bp mRNA linear EST 11-OCT-2001
 LOCUS 603391922F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5402236 5',
 DEFINITION mRNA sequence.
 ACCESSION BI688773
 VERSION BI688773.1 GI:16042446
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 768)
 AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM12025 row: k column: 05
 High quality sequence start: 2
 High quality sequence stop: 761.

FEATURES
 Location/Qualifiers
 1..768
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5402236"
 /clone_lib="NIH_MGC_90"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.7 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."
 BASE COUNT 271 a 177 c 164 g 156 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 2 14e-97 Length: 768
 Score: 1140.00 Matches: 243
 Percent Similarity: 95.72% Conservative: 3
 Best Local Similarity: 94.55% Mismatches: 10
 Query Match: 37.82% Indels: 5
 DB: 13 Gaps: 0

US-09-830-144-2 (1-579) x B1668773 (1-768)

QY 315 SerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsn 334
 DB 3 AGTACAGGCTCATTCATGACATCTTCTTACAAATACAGATACAAAGTGACACTAAT 62
 QY 335 MetGluGluValProIleThrAsnAspThrIleValArgLeuGluSerLysLeuLys 354
 DB 63 ATGGACCAAGTTCTCTCCCAAAATGATTAATTAAGCGCTTAGAAATCAAAATG-TTGAAA 121
 QY 355 AsnGlnAlaLysGlnInserGlnSerGlyArgLeuSerLeuGlyAlaSerHisGlySer 374
 DB 122 AATCAGGCAAGCAACAGATGATCTGACGTTTAAGTTTGGAGCCCTCCGTGGAGC 181
 QY 375 SerValGluSerLeuProProThrSerGluGlyLysArgMetSerAlaAspMetSerGlu 394
 DB 182 AGGTGGAAGAGCTTGGCCCCAACCTGTAGGGCAAGAGATGATGCTGACATGCTCGAA 241
 QY 395 IleGluAlaArgIleAlaIleThrThrArgLysAsnGlyLysProArgArgSerIleGln 414
 DB 242 ATGAAAGCTAGAGTCCGCCCAACACAGGCAACGACACGACAGATGATCCATCCAA 301
 QY 415 AspLeuThrValThrGlyThrGluProGlyGlnValSerSerArgSerSerProSer 434
 DB 302 GACTTGACTGTAAGTGAACAGACCTGTCAGGTGAGCAGTATGATCCAGTCCAGT 361
 QY 435 ValArgMetIleThrThrSerGlyProThrSerGluLysProThrArgSerHisProThr 454
 DB 362 GTCAGATGATTAATTAATCCTCAGGACCAACCTCAGAAAGCCAACTCGAAGTATCATG 421
 QY 455 ThrProAspAspSerThrAspThrAsnGlySerAspAspSerIleProMetAlaIleLeu 474
 DB 422 ACCCTGATGATTTCCACAGATACCAATGATCAATTAATCCATCCCAATGGCTTATCTT 481
 QY 475 ThrLeuAspHisGlnLeuGlnProLeuAlaProCysProAsnSerLysGluSerMetAla 494
 DB 482 ACACGTGATCAACCACTACAGCCTCTAGACCGTCCCAAACTCCAAAGATCTATGCA 541
 QY 495 ValPheGluGlnHisCysLysMetAlaGlnGlyIleMetLysValGlnThrGluIleAla 514
 DB 542 GGTGTTGAACAGCATTTGAAATGGCACAGATATATGAAAGTTCAAAACGAAATGAT 601
 QY 515 LeuLeuLeuGlnArgLysGlnLeuValAla-GluLeuAspGlnAspGluLysAspGln 534
 DB 602 CGTTATTACAGAGAAACAGAACTAGTTGACAGAACTGACCGAGATGAAAGGACCA 661
 QY 534 GlnAsnThrSerArgLeuValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLe 554
 DB 662 GCAAAATACATCTCGCTGCTGACAGAACTTAAAGCTTTAGATG-AAAAACAAGACT 720
 QY 554 SerThrIleThrArgGlnGlnCysLysLeuGlnLeuGluValIleArgSer 570
 DB 721 TTCTACTTAACACAGCA-TGCMAAAACA-CTAGAGGTCACTCAGAACT 767
 RESULT 7
 AL550589 827 bp mRNA linear EST 16-FEB-2001
 LOCUS AL550589 LTI_NFL006.PL2 Homo sapiens cDNA clone CS01058Y014.5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL550589
 VERSION AL550589.1 GI:12887700
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 827)
 AUTHORS Li, W. B., Gruber, C., Jesse, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France

FEATURES
 source
 Email: segref@genoscope.cns.fr, Web: www.genoscope.cns.fr.
 Location/Qualifiers
 1..827
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS01058Y014"
 /clone_lib="LTI_NFL006_PL2"
 /tissue_type="placenta"
 /note="Vector: pcwvSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT)-primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pcwvSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, Rockville, Maryland 20850, USA Fax: (1) 301 610 8371
 Email: fliang@life.com URL: <http://fulllength.invitrogen.com>

BASE COUNT 196 a 178 c 228 g 209 t 16 others
 ORIGIN

Alignment Scores:
 Pred. No.: 2,33e-95 Length: 827
 Score: 1119.00 Matches: 216
 Percent Similarity: 97.32% Conservative: 2
 Best Local Similarity: 96.43% Mismatches: 5
 Query Match: 37.13% Indels: 1
 DB: 9 Gaps: 0

US-09-830-144-2 (1-579) x AL550589 (1-827)

QY 1 MetSerThrAlaSerAlaLysSerSerSerSerSerAlaGlyGluMetIleGlu 20
 DB 157 ATGCTTAAAGCCTCTGCGGCTCTCMTCTCTGCTTGGCGGCTTAAATGATCGAA 216
 QY 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
 DB 217 GCCCTTCCAGGTCCTCACTTGAAGATCGATCAACAGAGATCAAGGTGGAAAG 276
 QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
 DB 277 GTTGTTGAAGAGGAGCCTTGGARTTTTGGCAAGCTAGATGGAGCAAAATATGTT 336
 QY 61 AlaIleLysGlnIleGlnSerGlnSerGluArgLysAlaPheIleValGluLeuArgGln 80
 DB 337 GCTATTAAATAAATAAAGTGAATCTGAGAGAAAGCTTATTATGAGCTTGGGAG 396
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
 DB 397 TTATCCCGTGAACCATCTTAATATGTAAGCTTTATGAGCGCTGCTGAATCCAGTG 456
 QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyValGlu 120
 DB 457 TGTCTTGATGAATATCTGAAGGGGCTCTTATATTAATGTCATGCTGCTGAA 516
 QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln-Gl 140
 DB 517 CCATTCCCATATTATCTCTGCTCCACAGCAATGATGATGTTTACAGGTGCCAAGG 576
 QY 140 yValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPr 160
 DB 577 AGTGCTTATCTTCAGAGATGCAACCCAAAGCGCTTAATTCACAGGAGCTGMAACAC 636
 QY 160 oAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
 DB 637 AAACCTACGCTGTTGGACGGGGAGCACTTAAATAATTTGTGTTGGTACGCTG 696
 QY 180 sAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
 DB 697 TGACATTCAAGACACATGACCAATTAACAAGGGAGTCTGCTGATGACCACTGAAGT 756
 QY 200 lPheGluGlySerAsnTyrSerGluLysCysAspValAlaHisSerTrpGlyIleIleLeuTr 220


```

Db 757 TTTTGAAGGTAGTAATTACAGTCGAAAAAATGTGACGTCTTCAGCTGGGTATTATTCTTTG 816
Qy 220 pgluValile 223
Db 817 GGAAGTGATA 826

RESULT 8
BJ062988 688 bp mRNA linear EST 10-DEC-2001
LOCUS BJ062988
DEFINITION laevis cDNA clone XL069m17 5', mRNA sequence.
ACCESSION BJ062988
VERSION BJ062988.1 GI:17470746
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 688)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1. .688
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/notes="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 206 a 139 c 154 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3.33e-95 Length: 688
Score: 1116.00 Matches: 204
Percent Similarity: 96.31% Conservative: 5
Best Local Similarity: 94.01% Mismatches: 8
Query Match: 37.03% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2 (1-579) x BJ062988 (1-688)

Qy 120 GluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 139
Db 22 GAACCTTTGCTTACTATCTACTGCTGCCCATGCAATGAGTTGGTGTTCACAAATGTGCCCAA 81
Qy 140 GlyValAlaTyrLeuHisSerMetGlnProLeuAlaLeuIleHisArgAspLeuLysPro 159
Db 82 GGAGTTGCATATTTACATGATGAGCCAAAGCGCTCTGATTCACAGGACCTCAAAOCA 141
Qy 160 ProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 179
Db 142 CCAAACTGTTGTGTAGCTGAGGACACTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 180 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 199
Db 202 TGTGATATTACAGCTCACATGACTAATAACAAAGGAAGTGCAGCATGGATGGCTCCAGAA 261

```

```

Qy 200 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleIeu 219
Db 262 GTTTTGAAGGTAGCACTACAGCGAAAAATGTGACGTGTTAGTTGGGGCATTTATCTT 321
Qy 220 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgile 239
Db 322 TGGGAAGTAATAACCGAAGAAAAACCTTTTCGATGAAATTGGTGGTCCAGCGTTCGGTATA 381
Qy 240 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259
Db 382 ATGTGGGCTGTTCAATGTGCTACTCGGCCACCATTAATAATAAAAAATTTGCCCTAACCCATT 441
Qy 260 GluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGlu 279
Db 442 GAAAGCTTAATGATCTCGCTGCTGTCCAAAGATCCCCACAAAACCTTCAATGGAGGAG 501
Qy 280 IleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 299
Db 502 ATTGTCAAGATAATGACACATCTAAAGCAGTATTTTCTGTGAGCAGACGTTTCCTTACAG 561
Qy 300 TyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPhe 319
Db 562 TATCCTTGTCTAGTACTCTGATGAAGGGCAAGCAATTTCTGCACAAAGTACAGGTTCTCTGT 621
Qy 320 MetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGlu 336
Db 622 ACTGATATTACATCTACTAATACTAGTAACAAGAGTGAAGTCAATATAGAA 672

RESULT 9
BJ074867 696 bp mRNA linear EST 11-DEC-2001
LOCUS BJ074867
DEFINITION laevis cDNA clone XL071110 5', mRNA sequence.
ACCESSION BJ074867
VERSION BJ074867.1 GI:17505056
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 696)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1. .696
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/notes="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT 210 a 142 c 155 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3.39e-95 Length: 696

```

Score: 1116.00 Matches: 204
Percent Similarity: 96.31% Conservative: 5
Best Local Similarity: 94.01% Mismatches: 8
Query Match: 37.03% Indels: 0
Gaps: 13

US-09-830-144-2 (1-579) x BU074867 (1-696)

```
QY 120 GUpProleuBrOTyTyrThraAlaAlaHisAlaMetSerTpcCysLeuGncCysSerGln 139
Db 22 GAACCTTGCTTACTACTACTGCTGCCCATGCAATGAGTGTGTTCATCAATGTGCCAA 81
QY 140 GilyAlaAlaYrleuHsSerMetGlnProlyAlaLeuIleHisArgAspLeuPro 159
Db 82 GGAGTTCACATTTTACATACATGAAAGCCAAAGCTCTGATTCACAGGACCTCAACCA 141
QY 160 ProAsnLeuLeuValaIaGlyGlyThrValleuYsIleCysAspPheGlyThra 179
Db 142 CCAAACTTGCTGCTGAGCTGAGGACCTGTTCTTAAGATTGTGACTTGTGACAGCC 201
QY 180 CysAspIleGlnThrHisMetThraAsnAsnlySglySerAlaAlaTrrPheAlaProGlu 199
Db 202 TGCGATATTTCAGACTCAGCATGACTTAATTAACAAGAGTGCAGCATGATGCTCCAGAA 261
QY 200 ValPheGlySerAsnTyrSerGlyLysCysAspValPheSerTrrGlyIleIleLeu 219
Db 262 GTTTTGAAGTACACATACAGCAAGAAAATGACGTGTTTGAAGGCGCATTTCTT 321
QY 220 ThrGlyValIleThraArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 239
Db 322 TGGGAATTAATTAACCCGAAAGAAACCTTTGATGAATTTGGTGCAGGCTTCCATA 381
QY 240 MetTrrAlaValHisAsnGlyThraArgProProleuIleLysAsnLeuProLysProIle 259
Db 382 ATGGGGCTGCTTCCATCAATGATGCTCGGCCCATTAATTAATAATTGCTTAAGCTATT 441
QY 260 GluSerLeuMetThraGlyCSTrrPserLysAspProSerGlnArgProSerMetGluGlu 279
Db 442 GAAAGCTTAATGACTCGCTCTGCTGTCGAAAGATCCCCCAAAAGACCTTCAATGAGGAG 501
QY 280 IleValIysIleMetThrHisIleuMetArgTyrPheProGlyValaAspGluProLeuGln 299
Db 502 ATGTCTAAGTAAATGACACATCTTAAGACATATTTCTTGAGACACGCTTCTTACAG 561
QY 300 TyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrglySerPhe 319
Db 562 TATCTCTTGATAGTACTGTATGAAGGCAAGCATTTCTGCACAAAGTACAGGTTCTGT 621
QY 320 MetAspIleAlaSerThraAsnThrSerAsnLysSerAspThrasMetGlu 336
Db 622 ACTGATATTACATCTACTAATACTAGTAAACAGATGACATCAATATAGAA 672
RESULT 10
LOCUS BQ219348 1006 bp mRNA linear EST 02-MAY-2002
DEFINITION AGENCOURT_7260864 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785294
ACCESSION BQ219348
VERSION BQ219348.1 GI:20400748
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE NIH-MGC http://mgs.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
```

DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL2872 row: c column: 23
High quality sequence stop: 624.

FEATURES

SOURCE

```
1..1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5785294"
/clone_lib="NIH MGC 71"
/tissue_type="leiomysarcoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pcwv-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 240 a 240 c 293 g 233 t
ORIGIN
```

Alignment Scores:

Pred. NO.: 6,156-95 Length: 1006
Score: 1116.00 Matches: 226
Percent Similarity: 84.31% Conservative: 5
Best Local Similarity: 82.48% Mismatches: 17
Query Match: 37.03% Indels: 27
DB: 14 Gaps: 4

US-09-830-144-2 (1-579) x BQ219348 (1-1006)

```
QY 1 MetSerThraAlaSerAlaAlaSerSerSerSerSerSerAlaGlyLumetIleGlu 20
Db 163 ATGTCTAAGCTTGCTGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 222
QY 21 AlaProSerGlnValleuAsnPhelGluIleAspTyrHsGlyIleGluValaGluGlu 40
Db 223 GCCCTTCCCGAGTCTCTCACTTTGAAGAGATCCACACAGAGATCGAGTGGAAAG 282
QY 41 ValValGlyArgGlyAlaPheGlyValaValaCysLysAlaIleSTrrPargAlaLysAspVal 60
Db 283 GTTGTGGAAGAGAGAGCTTGTGAGTTGTTGCAAACTCAAGTGAAGCAAAAGATGT 342
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValaGluLeuArgLys 80
Db 343 GCTATTAAACAAATAGAAAGGATCTGAGAGAAACGTTTATTATGAGCTTCGGCAG 402
QY 81 LeuSerArgValaAsnHisProAsnIleValLysLeuTrrGlyAlaCysLeuAsnProVal 100
Db 403 TTATCCCGTGGAAACCATCTTAATATTGTAAAGCTTTATGAGCGCTGGAATCCAGTG 462
QY 101 CysLeuValMetGluTrrAlaGlyIleGlySerLeuTrrAsnValleuHsGlyAlaGlu 120
Db 463 TGTCTTGATGGAATATGCTGGAAGGGGGCTTTATATATATAGTGCATGGGTGAA 522
QY 121 ProLeuProTrrTrrThraAlaAlaHisAlaMetSerTrrCysLeuGncCysSerGlnGly 140
Db 523 CCATTGCCATATTATACGCTGCCACGACATGATGTTGGTTCAGTGTTCACAGGA 582
QY 141 ValAlaTrrLeuHsSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db 583 GTGGCTTATCTTCCACAGCATGCAACCAAGCGTTAATTCACAGGACCTGAAACCA 642
QY 161 AsnLeuLeuValaIaGlyGlyThrValleuYsIleCysAspPheGlyThraAlaCys 180
Db 643 AACTTACGCTGCTGTTGACAGGGGGACAGTTCT-AAAATTTGTGATTTGGGACAGCTGT 701
QY 181 AspIleGlnThrHisMetThraAsnLysGlySerAlaAlaTrrPheAlaProGluVal 200
Db 702 GGACTTCAGACACATGACCAATTAACAAGGGGCTCTCTGGAAGGACCTCGAATT 761
QY 201 PheGlyLysSerAsnTrrSerGlyLysCysAspValPheSerTrrGlyIleIleLeuTrr 220
Db 762 TTTGGAGGAGTAATTAACCGGAAAAAATGACGCTTTCAGCTGGGGAATTTTCTTTGG 821
```


QY 420 G|YThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThr 439
DB 541 GGGACGAGAACCTGTCAGAGGACCGGTCATCCACCTAGTGTGAGAAATGATCACT 600
QY 440 ThrSerGlyProThrSerGlyValProThrArgSerHisProThrProAspSer 459
DB 601 ACCTCAAGACCACTCAAGAAAGCACTCGCATCACCCTGGACCCCTGATGATTC 660
QY 460 Thr 460
DB 661 ACA 663
RESULT 12 646 bp mRNA linear EST 11-DEC-2001
LOCUS BJ073883
DEFINITION BJ073883 NIBB Mochii normalized Xenopus tailbud library Xenopus laevis cDNA clone XLI01n05 5', mRNA sequence.
ACCESSION BJ073883 GI:17504072
VERSION BJ073883.1
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 646)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadau Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES
source
1..646
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="XLI01n05"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/cissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 190 a 132 c 143 g 181 t
ORIGIN

Alignment Scores:

Pred. No.: 2,07e-92 Length: 646
Score: 1086.00 Matches: 203
Percent Similarity: 96.26% Conservative: 3
Best Local Similarity: 94.86% Mismatches: 8
Query Match: 36.03% Indels: 1
Gaps: 0

US-09-830-144-2 (1-579) x BJ073883 (1-646)

QY 114 AsnValIeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyr 133
DB 4 AATGTTTGCATGAGCTGAACCTTGCTTACAT-ACGTGCTCCCATGCAATGAGTTGG 62
QY 134 CysIeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaIleu 153
DB 63 TGTTTACAAATGTCGCCAAGAGTTGCAATTTTACATAGCATGAAGCAAGGCTTGATT 122

QY 154 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 173
DB 123 CACAGGAGCCTCAACCAACCAACTTGTCTGTGAGCTGAGGCACTGTTCTTAAGATT 182
QY 174 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 193
DB 183 TGTGACTTTGGTACAGCTGTGATATTACAGCTCAGCTGACTATATAACAAGAAAGTGCA 242
QY 194 AlaTrpMetAlaProGluValPheGlyGlySerAsnTyrSerGlyLysCysAspValPhe 213
DB 243 GCATGATGTGCTCCAGAAAGTTTGAAGTACCACTACAGCCAAAATGTGACGTGTTT 302
QY 214 SerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 233
DB 303 AGTTGGGCAATTATCTTTGGGAAAGTAATACCGAAGAAACCTTTCATGAAATTTGGT 352
QY 234 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 363 GGTCCAGGCTTCCTATATATGTGGCTGTTCACATGTGTACTGGCCACATTATTAATA 422
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
DB 423 AATTGGCTTAAGCTATTGAAGCTTATGACTGCTGCTGTCACAAAGATCCCCACAA 482
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
DB 483 AGACCTTCATGAGGAGGATTTGTAAGATATGACATCTTAAGCAAGTATTTTCCCGGA 542
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
DB 543 GCAGACGTTCTTACAGTATCTCTTGTCAGTACTGATATGAAGGCAACCAATTCTGCA 602
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThr 327
DB 603 ACAAGTACAGGTCCTGACTGATATTACATCTACTACT 644

RESULT 13 604 bp mRNA linear EST 30-JAN-2002
LOCUS BM426610
DEFINITION pgf2n.pk004.d15 Normalized Chicken Abdominal Fat Library (pgf2n)
Gallus gallus cDNA clone pgf2n.pk004.d15 5' similar to
emb|CAB87605.1 (Alu21964).d154614.1.4 (mitogen-activated protein
kinase kinase kinase 7 (TGF-beta activated kinase id (TRAK1))) [Homo
sapiens], mRNA sequence.
ACCESSION BM426610
VERSION BM426610.1 GI:18431087
KEYWORDS EST.
SOURCE chicken.
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 604)
AUTHORS Cogburn,L.A., Morgan,R. and Burnside,J.
ESTs from Normalized Chicken fat cDNA library-USDA/IFARS Animal
Genome Project
JOURNAL Unpublished (2002)
COMMENT Contact: Larry A. Cogburn
University of Delaware
Townsend Hall, Newark, DE 19717, USA
Tel: 302-831-1335
Fax: 302-831-2822
Email: cogburn@udel.edu, www.chickest.udel.edu.

FEATURES
source
1..604
/organism="Gallus gallus"
/strain="Commercial broiler, Ottawa Research Centre,
leghorn"
/db_xref="taxon:9031"
/clone="pgf2n.pk004.d15"
/clone_lib="Normalized Chicken Abdominal Fat Library
(pgf2n)"
/sex="Male and Female"

```

/tissue_type="Abdominal Fat"
/dev_stage="Embryonic (d18,d19) ; post-hatch (d1,w3,w7,w9,w16,1yr)"
/lab_host="E. coli EMDH10B"
/note="Vector: pCMVSPORT6; Library made from equivalent pools of total RNA isolated from each developmental age (across strains) ; Single pass sequencing from 5'-end"
167 a 125 c 140 g 167 t
BASE COUNT

```

Alignment Scores:		
Pred. No.:	8.52e-92	Length: 604
Score:	1079.00	Matches: 196
Percent Similarity:	98.51%	Conservative: 2
Best Local Similarity:	97.51%	Mismatches: 3
Query Match:	35.80%	Indels: 0
DB:	13	Gaps: 0

US-09-830-144-2 (1-579) x BM426610 (1-604)

Qy	105	GluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGluProLeuProTyr	124
Db	2	GAGTATGCTGAAGGAGTTCTCTGTACAATGTCTGATGGTCTGAACCTCTGCCTCAT	61
Qy	125	TyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeu	144
Db	62	TATACTGTCGACACGCAATGAGTTGGTGTATTACAGTGTTCCTCAAGGAGTGGCATATCTT	121
Qy	145	HisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProProAsnLeuLeuLeu	164

165	ValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr	184
Qy		
182	GTACTGGGGGACAGTCTTAAGATCTGTGATTTTGGTACACCTGTGATATTCAACA	241
Db		
185	HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer	204
Qy		
242	CACATGACCAACAATAAGGGAAGTCTGCTTTGGATGGCACTGAAGTTTTGAAGGTAC	301
Db		
205	AsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr	224
Qy		

225	ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHis	244
Qy		
362	CGTAGGAAACCTTTTGATGAGATTGGTGGTCCAGCTTCGCGATAATGTGGCGAGTTCCAC	421
Db		
245	AsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThr	264
Qy		
422	AATGGTACTCGACACACACCTGATCAAAACCTTACCTAAACCAATTCGAGAGTTTAAATGACC	481
Db		

Qy	265	ArgCysTrpSerIysAspProSerCysArgProSerMetCysGluLeuValIysIleMet	284
Db	482	CGCTGTGGTCCAGGATCCCTCAACAGCACCTCCAGGAGAAATGTTAAAAATATG	541
Qy	285	ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyr	304
Db	542	ACACACTTGATGCGGTACTTTCCNNNAGCTGATGAACCTCTGCAGTATCCTTCNAGTAT	601

Qy	305	Ser	305
D ^b	602	TCA	604

[illegible]

ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 929)
Li.W.B., Gruber.C., Jesse,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr

FEATURES	SOURCE
1. <i>Age</i>	1. <i>Age</i>
2. <i>Gender</i>	2. <i>Gender</i>
3. <i>Marital Status</i>	3. <i>Marital Status</i>
4. <i>Education</i>	4. <i>Education</i>
5. <i>Income</i>	5. <i>Income</i>
6. <i>Occupation</i>	6. <i>Occupation</i>
7. <i>Religion</i>	7. <i>Religion</i>
8. <i>Political Affiliation</i>	8. <i>Political Affiliation</i>
9. <i>Health Status</i>	9. <i>Health Status</i>
10. <i>Travel History</i>	10. <i>Travel History</i>
11. <i>Employment Status</i>	11. <i>Employment Status</i>
12. <i>Family Size</i>	12. <i>Family Size</i>
13. <i>Home Ownership</i>	13. <i>Home Ownership</i>
14. <i>Vehicle Ownership</i>	14. <i>Vehicle Ownership</i>
15. <i>Insurance Status</i>	15. <i>Insurance Status</i>
16. <i>Subscription Services</i>	16. <i>Subscription Services</i>
17. <i>Charitable Donations</i>	17. <i>Charitable Donations</i>
18. <i>Volunteer Work</i>	18. <i>Volunteer Work</i>
19. <i>Real Estate Transactions</i>	19. <i>Real Estate Transactions</i>
20. <i>Banking History</i>	20. <i>Banking History</i>
21. <i>Investment Portfolio</i>	21. <i>Investment Portfolio</i>
22. <i>Travel Preferences</i>	22. <i>Travel Preferences</i>
23. <i>Food and Dining Habits</i>	23. <i>Food and Dining Habits</i>
24. <i>Shopping Behavior</i>	24. <i>Shopping Behavior</i>
25. <i>Healthcare Usage</i>	25. <i>Healthcare Usage</i>
26. <i>Education Enrollment</i>	26. <i>Education Enrollment</i>
27. <i>Professional Certifications</i>	27. <i>Professional Certifications</i>
28. <i>Language Proficiency</i>	28. <i>Language Proficiency</i>
29. <i>Artistic Interests</i>	29. <i>Artistic Interests</i>
30. <i>Sports Participation</i>	30. <i>Sports Participation</i>
31. <i>Technology Usage</i>	31. <i>Technology Usage</i>
32. <i>Environmental Concerns</i>	32. <i>Environmental Concerns</i>
33. <i>Religious Observance</i>	33. <i>Religious Observance</i>
34. <i>Political Engagement</i>	34. <i>Political Engagement</i>
35. <i>Healthcare Preferences</i>	35. <i>Healthcare Preferences</i>
36. <i>Education Goals</i>	36. <i>Education Goals</i>
37. <i>Professional Aspirations</i>	37. <i>Professional Aspirations</i>
38. <i>Language Learning</i>	38. <i>Language Learning</i>
39. <i>Artistic Skills</i>	39. <i>Artistic Skills</i>
40. <i>Sports Interests</i>	40. <i>Sports Interests</i>
41. <i>Technology Skills</i>	41. <i>Technology Skills</i>
42. <i>Environmental Awareness</i>	42. <i>Environmental Awareness</i>
43. <i>Religious Practices</i>	43. <i>Religious Practices</i>
44. <i>Political Views</i>	44. <i>Political Views</i>
45. <i>Healthcare Choices</i>	45. <i>Healthcare Choices</i>
46. <i>Education Plans</i>	46. <i>Education Plans</i>
47. <i>Professional Development</i>	47. <i>Professional Development</i>
48. <i>Language Acquisition</i>	48. <i>Language Acquisition</i>
49. <i>Artistic Hobbies</i>	49. <i>Artistic Hobbies</i>
50. <i>Sports Activities</i>	50. <i>Sports Activities</i>
51. <i>Technology Proficiency</i>	51. <i>Technology Proficiency</i>
52. <i>Environmental Involvement</i>	52. <i>Environmental Involvement</i>
53. <i>Religious Beliefs</i>	53. <i>Religious Beliefs</i>
54. <i>Political Participation</i>	54. <i>Political Participation</i>
55. <i>Healthcare Decisions</i>	55. <i>Healthcare Decisions</i>
56. <i>Education Outcomes</i>	56. <i>Education Outcomes</i>
57. <i>Professional Achievements</i>	57. <i>Professional Achievements</i>
58. <i>Language Fluency</i>	58. <i>Language Fluency</i>
59. <i>Artistic Talent</i>	59. <i>Artistic Talent</i>
60. <i>Sports Skills</i>	60. <i>Sports Skills</i>
61. <i>Technology Expertise</i>	61. <i>Technology Expertise</i>
62. <i>Environmental Activism</i>	62. <i>Environmental Activism</i>
63. <i>Religious Devotion</i>	63. <i>Religious Devotion</i>
64. <i>Political Involvement</i>	64. <i>Political Involvement</i>
65. <i>Healthcare Choices</i>	65. <i>Healthcare Choices</i>
66. <i>Education Attainment</i>	66. <i>Education Attainment</i>
67. <i>Professional Growth</i>	67. <i>Professional Growth</i>
68. <i>Language Proficiency</i>	68. <i>Language Proficiency</i>
69. <i>Artistic Interests</i>	69. <i>Artistic Interests</i>
70. <i>Sports Participation</i>	70. <i>Sports Participation</i>
71. <i>Technology Usage</i>	71. <i>Technology Usage</i>
72. <i>Environmental Concerns</i>	72. <i>Environmental Concerns</i>
73. <i>Religious Observance</i>	73. <i>Religious Observance</i>
74. <i>Political Engagement</i>	74. <i>Political Engagement</i>
75. <i>Healthcare Preferences</i>	75. <i>Healthcare Preferences</i>
76. <i>Education Goals</i>	76. <i>Education Goals</i>
77. <i>Professional Aspirations</i>	77. <i>Professional Aspirations</i>
78. <i>Language Learning</i>	78. <i>Language Learning</i>
79. <i>Artistic Skills</i>	79. <i>Artistic Skills</i>
80. <i>Sports Interests</i>	80. <i>Sports Interests</i>
81. <i>Technology Skills</i>	81. <i>Technology Skills</i>
82. <i>Environmental Awareness</i>	82. <i>Environmental Awareness</i>
83. <i>Religious Practices</i>	83. <i>Religious Practices</i>
84. <i>Political Views</i>	84. <i>Political Views</i>
85. <i>Healthcare Choices</i>	85. <i>Healthcare Choices</i>
86. <i>Education Plans</i>	86. <i>Education Plans</i>
87. <i>Professional Development</i>	87. <i>Professional Development</i>
88. <i>Language Acquisition</i>	88. <i>Language Acquisition</i>
89. <i>Artistic Hobbies</i>	89. <i>Artistic Hobbies</i>
90. <i>Sports Activities</i>	90. <i>Sports Activities</i>
91. <i>Technology Proficiency</i>	91. <i>Technology Proficiency</i>
92. <i>Environmental Involvement</i>	92. <i>Environmental Involvement</i>
93. <i>Religious Beliefs</i>	93. <i>Religious Beliefs</i>
94. <i>Political Participation</i>	94. <i>Political Participation</i>
95. <i>Healthcare Decisions</i>	95. <i>Healthcare Decisions</i>
96. <i>Education Outcomes</i>	96. <i>Education Outcomes</i>
97. <i>Professional Achievements</i>	97. <i>Professional Achievements</i>
98. <i>Language Fluency</i>	98. <i>Language Fluency</i>
99. <i>Artistic Talent</i>	99. <i>Artistic Talent</i>
100. <i>Sports Skills</i>	100. <i>Sports Skills</i>

```
/organism="Homo sapiens"  
/db_xref="taxon:9606"  
/clone="CSDB002YM17"  
/clone_lib="LTI_NFL004_NBC2"  
/sex="male"  
/tissue_type="neuroblastoma cells"  
lab_host="IMH10B"
```

/lab_host="DHUBS"
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed
 by Life Technologies. Contact : Feng Liang Life
 Technologies, a division of Invitrogen 9800 Medical Center
 Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
 8371 Email : fliang@lifetech.com URL :
 http://www.lifetech.com"

BASE COUNT	213 a	235 c	249 g	222 t	10 others
ORIGIN	http://fulllength.invitrogen.com				

Alignment Scores:	
Pred. No.:	5,586-90
Score:	1063.00
Percent Similarity:	100.00%
Best Local Similarity:	99.51%
Query Match:	35.27%
DB:	9
Gaps:	0
Indels:	0
Mismatches:	0
Conservative:	1
Matches:	204
Length:	929

US-09-830-144-2 (1-579) x AL520975 (1-929)

[illegible]

DB 725 GTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGGAGCCTGAACCCACCA 784
QY 161 AsnleuleuleuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCys 180
DB 785 AACCTACTGCTGTGGAGGGGAGCAGTTCTAAATTTGTGATTTTGTGACAGCCTGT 844
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
DB 845 GACATTCAGACACATGACCAATAACAGGGAGTGTGCTTGATGGCMCTGAAGTT 904
QY 201 PheGluGlySerAsn 205
DB 905 TTGGAAGCTAGTAAT 919
RESULT 15
BG548917 686 bp mRNA linear EST 04-APR-2001
LOCUS 602575750F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4703580 5',
DEFINITION mRNA sequence.
ACCESSION BG548917 GI:13547582
VERSION BG548917.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 686)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bbs-r@mail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM1542 row: d column: 13
High quality sequence stop: 671.
Location/Qualifiers
FEATURES
source 1..686
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703580"
/clone_1ib="NIH_MGC_77"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: Lung; Vector: pDNR-LIB (Clontech); Site: 1:
SfiI (ggccgctcgcc); Site_2: SfiI (ggccatctatggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTAATGGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGGAGCGGCGGCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."

BASE COUNT 202 a 143 c 169 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 5,28e-90 Length: 686
Score: 1061.00 Matches: 215
Percent Similarity: 94.78% Conservative: 3
Best Local Similarity: 93.48% Mismatches: 8
Query Match: 35,20% Indels: 5
DB: 12 Gaps: 0

US-09-830-144-2 (1-579) x BG548917 (1-686)
QY 155 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValIleuLysIleCys 174
|||||

DB 4 AGGAGCTGAACACCAACCAACTTACTCTGCTAGCTGGGGGACAGTTCTTAAAAATTTGT 63
QY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
DB 64 GATTTGGTACAGCCTGTGACATTGACACACATGACCAATTAACAAGGAGAGTGCTGT 123
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu-LysCysAspValPhe 214
DB 124 TGATGGACCTGAAGATTTTGAAGTAGTAATTACAGTAACCAATGATGACCTCTTCAG 183
QY 214 TTPGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 224
DB 184 CTGGGGTATTATTCTTTGGGAAGTGAATACCGCTGGAAAACTTTGATAGATTGGTGG 243
QY 224 YProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProPheIleLys-A 254
DB 244 CCCAGCTTCCGAATCATGTGGGCTGTTCAATAATGTAATCGNACAGCACTGATTAACAA 303
QY 254 snLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln- 273
DB 304 ATTTACCTTAAGCCCATTTGAGAGCCTGATGACTGCTAGTGTCTTAAAGATCTTCCAGG 363
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
DB 364 CGCCCTTCATAGAGAGCAAAATTTGTAAATTAATGACTCACTTGATCGGTTCTTCCAGGA 423
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
DB 424 GCAGATGAGCCATTACAGTATCTCTTGACAGTATTCAGATGAGACAGACAACTCTGCC 483
QY 314 ThrSerThrGlySerPheMet-AspIleAlaSerThrAsnThrSerAsnLysSerAspTh 333
DB 484 ACCAGTACAGGCTCATTCATGAGCACATTTGCTTCAAAAATACAGATTAACAAAAGTGACAC 543
QY 333 rAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeu 353
DB 544 TTAATATGAGCAAAATTTCTCTCCCAAAATGATCATTTAAACCGCTTAACATCAAC-ATG 602
QY 353 uLysAsnGlnAlaLysGlnGlnSerGlnSerGlyArgLeuSerLeuGlyAlaSerHisG 373
DB 603 GAACAAATCAGGCAACAGCAACAGAGTGAATCTGGAGCTTTAAGCTTGGAGACCTCCCGTG 662
QY 373 ySerSerValGluSerLeuPro 380
DB 663 GAGCAGTGTGGAGAGCTTGCCC 684

Search completed: December 10, 2002, 02:03:51
Job time : 1821.11 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 23:00:44 ; Search time 61.482 Seconds
(without alignments)
2888.095 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKLEIVRSQQKRGQTS 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141354_2283/app.query.fasta.1.1422
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=Blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG 1.1.44 @runat_04122002_141354_2283 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT_LONGLOG -DEV TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCOTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	1788	4	US-09-529-279-14 Sequence 14, Appl
2	3014	100.0	2656	2	US-08-685-625A-5 Sequence 5, Appl
3	3014	100.0	2656	4	US-09-529-279-3 Sequence 3, Appl
4	2975	98.7	2433	2	US-08-685-625A-1 Sequence 1, Appl
5	490	16.3	2120	3	US-09-221-235-4 Sequence 4, Appl
6	490	16.3	2120	3	US-09-221-928-4 Sequence 4, Appl
7	490	16.3	2120	3	US-09-221-527-4 Sequence 4, Appl
8	490	16.3	2120	3	US-09-221-236-4 Sequence 4, Appl
9	490	16.3	2120	3	US-09-221-416-4 Sequence 4, Appl
10	490	16.3	2120	4	US-09-221-245-4 Sequence 4, Appl
11	490	16.3	2120	4	US-09-163-115-4 Sequence 4, Appl
12	490	16.3	2120	4	US-09-221-528-4 Sequence 4, Appl

13	490	16.3	2120	4	US-09-593-553-4	Sequence 4, Appl
14	490	16.3	2120	4	US-09-221-237-4	Sequence 4, Appl
15	481	16.0	1365	3	US-09-221-235-6	Sequence 6, Appl
16	481	16.0	1365	3	US-09-221-928-6	Sequence 6, Appl
17	481	16.0	1365	3	US-09-221-527-6	Sequence 6, Appl
18	481	16.0	1365	3	US-09-221-236-6	Sequence 6, Appl
19	481	16.0	1365	3	US-09-221-416-6	Sequence 6, Appl
20	481	16.0	1365	4	US-09-221-245-6	Sequence 6, Appl
21	481	16.0	1365	4	US-09-163-115-6	Sequence 6, Appl
22	481	16.0	1365	4	US-09-221-528-6	Sequence 6, Appl
23	481	16.0	1365	4	US-09-593-553-6	Sequence 6, Appl
24	481	16.0	1365	4	US-09-221-237-6	Sequence 6, Appl
25	418.5	13.9	3025	4	US-09-291-839-1	Sequence 1, Appl
26	418	13.9	2505	4	US-09-291-839-3	Sequence 3, Appl
27	412	13.7	3389	1	US-08-395-580-1	Sequence 1, Appl
28	412	13.7	3426	1	US-08-205-018-1	Sequence 1, Appl
29	398	13.2	2890	1	US-07-928-464-1	Sequence 1, Appl
30	398	13.2	2890	5	PCT-US93-07347-1	Sequence 1, Appl
31	398	13.2	3033	1	US-08-003-311B-1	Sequence 1, Appl
32	398	13.2	3033	1	US-08-261-432-1	Sequence 1, Appl
33	397	13.2	3516	3	US-09-188-930-257	Sequence 257, App
34	384	12.7	1888	3	US-09-188-930-66	Sequence 66, Appl
35	357.5	11.9	2574	4	US-09-142-529-2	Sequence 2, Appl
36	355.5	11.8	3132	1	US-08-167-919A-9	Sequence 9, Appl
37	355.5	11.8	3132	3	US-08-715-106-9	Sequence 9, Appl
38	354	11.7	4508	5	PCT-US93-06251-34	Sequence 34, Appl
39	349.5	11.6	3791	3	US-09-377-310-1	Sequence 1, Appl
40	346	11.5	3133	1	US-08-162-809-1	Sequence 1, Appl
41	345.5	11.5	2442	1	US-08-542-363-3	Sequence 3, Appl
42	345.5	11.5	2442	4	US-09-100-089-3	Sequence 3, Appl
43	345.5	11.5	2442	4	US-09-670-827-3	Sequence 3, Appl
44	345.5	11.5	6226	1	US-08-542-363-1	Sequence 1, Appl
45	345.5	11.5	6226	4	US-09-100-089-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-529-279-14
; Sequence 14, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TOSHIO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:
Pred. No.: 8.17e-272 Length: 1788
Score: 3014.00 Matches: 579
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-2 (1-579) x US-09-529-279-14 (1-1788)


```
Db 1803 GTACAGGAACATAAAAGCTTTAGATGAAACAAAAAGCCCTTTCTACTTACTACAGCAA 1862
Qy 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer 579
Db 1863 TGCAAAAAACAACATAGAGGTGCATCAGAGTCAGCAGCAGAAACGACAGGCACTTCA 1919

RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UENO, Naoto
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; TITLE OF INVENTION: TRANSDUCTION SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22131-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 157..1893
US-08-685-625A-1

Alignment Scores:
Pred. No.: 5,92e-268 Length: 2443
Score: 2975.00 Matches: 572
Percent Similarity: 98.79% Conservative: 0
Best Local Similarity: 98.79% Mismatches: 7
Query Match: 98.71% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-2 (1-579) x US-08-685-625A-1 (1-2443)

Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyLeuMetIleGlu 20
Db 157 ATGTGACAGCCCTCGCGCCCTCGTCCCTCTCTCTGTCGAGATGATCGAA 216
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
Db 217 GCGCGGTGCGAGTCTCTGAACCTTCGAGAGATCGACTACAGGAGATCGAGGTGGAAGAG 276
```

QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
 DB 1357 GCGACTGCAGGTAACGGGCAACCAAGGGGTAGATCATCCAAAGCTTACGTACTGG 1416
 QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr 440
 DB 1417 ACAGAACCTGGTCAGGTGAGCGCGGTATCCAGCCCTAGTGTCAGATGATCACTACC 1476
 QY 441 SerGlyProThrSerGlyLeuSerProThrArgSerHisProThrThrProAspSerThr 460
 DB 1477 TCAGGACCAACTCAGAGAGCCAGCTGCGATCCATCCATGAGTACCCCTGATCTCCACA 1536
 QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
 DB 1537 GACACCAATGGCTCAGATATACCTCCATCCAGTGGTATCTACACTGGATCACAGCTA 1596
 QY 481 GlnProLeuAlaProCysProAsnSerIleGlySerMetAlaValPheGlnGlnHisCys 500
 DB 1597 CAGCCCTCAGCGCGCGCCCAACTCCAAAGATCCATGGCAGTGTTCGAAACAGACGTGT 1656
 QY 501 LysMetAlaGlnGluTyrMetLeuValGlnThrGlnIleAlaLeuLeuLeuGlnArgLys 520
 DB 1657 AAATGTCACAGAGATATATGAAGTTCAACCGAAATCGCATTTGTTACTACAGAGAAAG 1716
 QY 521 GlnGluLeuValAlaGlnLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
 DB 1717 CAAGAACTGTTGACAAATTGGACCGAGATGAAGACCAAGCAAAATATCATCTCGTCTG 1776
 QY 541 ValGlnGlnHisLysLysLeuLeuAspGluLysSerLeuSerThrTyrTyrGlnGln 560
 DB 1777 GTACAGGAACATPAAAAAGCTTTTATGTAAGAAACAAAAGCCCTTTTACTATTAACAGCAA 1836
 QY 561 CysLysLysGlnLeuGlnValIleArgSerGlnGlnGlnLysArgGlnGlyThrSer 579
 DB 1837 TCGAAAAACAACCTAGAGGTCACTCAGAACCAACAGCAAGAGCAAGCAAGCACTTCA 1893

RESULT 5
 US-09-221-235-4
 : Sequence 4, Application US/09221235
 : Patent No. 6043040
 : GENERAL INFORMATION:
 : APPLICANT: Acton, Susan
 : TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 : FILE REFERENCE: MNI-050
 : CURRENT APPLICATION NUMBER: US/09/221,235
 : CURRENT FILING DATE: 1998-12-28
 : EARLIER APPLICATION NUMBER: 09/163,115
 : EARLIER FILING DATE:
 : NUMBER OF SEQ ID NOS: 15
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 4
 : LENGTH: 2120
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (47)..(1411)
 : US-09-221-235-4

Alignment Scores:
 Pred. No.: 4,97e-36 Length: 2120
 Score: 490.00 Matches: 138
 Percent Similarity: 48.79% Conservative: 84
 Best Local Similarity: 30.33% Mismatches: 159
 Query Match: 16.26% Indels: 74
 DB: 3 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-235-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GlnAlaProSerGlnVal 25
 DB 2 TCGACCAACGCGTCCGTTGAAAGTAAATCTTTGTCTATATGAGATGTGCTCTCCGGT 61

QY 26 LeuAsnPheGlnGluIleAspTyrLysGlnIleGluValGluGluValAlaGlyArgGly 45
 DB 62 GCCTCCCTTGTGCAAAATTAATTTGATGACTGCAAGTTTTTGAAGAACTGCGGTGAGGA 121
 QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
 DB 122 AGTTTGGAGTGTATTATGAGCCAAATGATATTCACAGACAGAGAGGTGCTGTAAAG 181
 QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 DB 182 AAGCTCTCAAAATAGAGAAAGAGCAGAA-----ATA 214
 QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro---- 99
 DB 215 CTGAGTCTCTCAGTCACAGAAACATCATCAAGTTTATGAGTAATTCTTGAACTCCC 274
 QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
 DB 275 AACTATGGCATTTGTACAGAAATATGCTTCTCTGGATCACTCTATGATTAACATTAACAGT 334
 QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
 DB 335 AACAGAGTGAGAGATG-----GATATGATCACAATTATGACTCTGGCCACTGAT 385
 QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
 DB 386 GTACCCAAAGAAAGCATTAATTATACATAGAGAGCTCTCTCAAGTGATTCACAGAGAC 445
 QY 157 LeuLysProProLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
 DB 446 CTCAAAGTCAGAAACGTTGTATAGCTGCGATGAG---GTACTGAAGATCTGTGACTTT 502
 QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
 DB 503 GGT---GCCTCTCGGTCCATTAACATACACACATGCTCTGTGTGAACCTTCCCA 559
 QY 195 TrpMetAlaProGluValAlaPheGlnGlySerAsnTyrSerGluLysCysAspValPheSer 214
 DB 560 TGGATGGCTCCAGAAAGTTATCCAGAGCTCTCTGTGCAAACTGTGACACATATTC 619
 QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 234
 DB 620 TATGTGTGTCTCTCGAGAGATGCTTAACAAGGAGGTCCCTTTAAAGTTTGAAGGA 679
 QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
 DB 680 -----TTACAAGTACGCTTGCTGTGAGTGAAGAAAAACAGAGATTAACTTCACAGC 733
 QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
 DB 734 AGTTGCCCCAGAAAGTTTGGCTGAACCTTTACATCAGTGTGGAGAGCTGATGCCAAGAAA 793
 QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
 DB 794 CGGCATCATTCAGCAAAATCATTTCAATCCTG----- 826
 QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
 DB 827 -----GAGTCCATGTCAAATGACAGC 847
 QY 314 ThrSerThrLysSerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 DB 848 -----AGCTTCTCGACAAAGTATCATTTCTTACACAAACAAAGCGGAGTGG 895
 QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
 DB 896 AGGTGCGAA---ATTGAAGCAACTCTTGAGAGCTTAAGAAACATGACCGTCACTCAGC 952
 QY 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
 DB 953 TTTAAGGACGACGAGCTTAAAGAAAGAAAGAGTTTAAAGATGTGGAGCAAAAGCTG 1012
 QY 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383

```

Db 1013 ACAGAGAGTCCAAACACCCCGCTTCTCTGCTGCTCAAGAAATGCTGAGGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluLeuGluAlaAaGlyAlaAa 401
Db 1073 TACTTTGAATCTAAACAGAGGAGTCAACAGTGCAGAGATGTCATGTCAGATCAACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAACGGGAGGCGCCATGGCATCAACCAAGTCTGCAGGCCATGCTGATGGGC 1192
QY 421 ThrGluProGlyGlnValSerArgSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

```

RESULT 6

```

US-09-221-928-4
; Sequence 4, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-928-4

```

Alignment Scores:

```

Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.28% Indels: 74
DB: 3 Gaps: 19

```

US-09-830-144-2 (1-579) x US-09-221-928-4 (1-2120)

```

QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACGCGTCCGGTGGGAAGTATAATCTTTGTCATTATGATGTCGTCTCTCGT 61
QY 26 LeuAsnPhaGluGluLeuAspTyrLysGluIleGluValGluGluValValGlyArgGly 45
Db 62 GCCTCTTTTGCAATTAATTAATTTGATGCTTGCAGTTTTTTTGAACCTGCGGTGGAGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGAGTGTATTATCGACCAATATGATATCACAGACACAGAGGTGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCTCAAAATAGAGAAGAGGCGAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro----- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTTATGGAGTAATCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGlyTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATGTTCACAGAAATATGCTTCTCTCGGATCACCTATGATTACATTAACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136

```

```

Db 335 AACAAAGTGGAGAGT-----CATATGATACATATTATGACCTGGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCAATTAATTTACATATGAGGCTCTGTCAAGGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAAAGTGTATTAGCTGCTGATGGA---GTACTCAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTCCATAACCATACACACATGTCCTTGGTTGGAACCTTCCCA 559
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAGTTATCCAGAGTCTCCCTGTGTGAGAACTTGTGACACATATTC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGTTGGTCTCTGGAGATGCTAACAGGAGGTCCCTTTAAAGGTTTGGAAGGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACGAGAGATTAAACATTCCAAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTTTGTGTAACCTGTATCATCAGTGTGGGAGCTGATGCCAAGAAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCATCATTTCAAGCAAAATCATTTCAATCCTG----- 826
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAAATGACACG 847
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTTCTCTGACAGTGTAACTTCATTCTACACAAACAGGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
Db 896 AGTGCAGAA---ATTGAGGCAACTCTTTGAGAGGCTAAAGAAACTAGAGCGTGTATCTCAGC 952
QY 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGAGCAGGAGCTTTAAGAACGAGAAAGACGTTTAAAGATGTGGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
Db 1013 ACAGAGCAGTCCAAACACCCCGCTTCTTTCCTCTTGTGCTGCAAGAATGTCTGAGGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla 401
Db 1073 TACTTTGAATCTAAACAGAGGAGTCAACAGTGCAGAGATGTCATGTCAGATCAACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAACGGGAGGCGCCATGGCATGAACCCCAAGTCTGCAGGCCATGATGCTGATGGGC 1192
QY 421 ThrGluProGlyGlnValSerArgSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

```

RESULT 7

```

US-09-221-527-4
; Sequence 4, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan

```


Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: 3 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-236-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
|||:::|||||:::|||||
Db 2 TCGACCCACGGCTCGGTGGAGTATATACTTTGTTCATTATGAGATGTCTCTCGGT 61

QY 26 LeuAsnPheGluIleAspTyrlLysGluIleGluValValGlyArgGly 45
|||:::|||||:::|||||
Db 62 GCCTCCTTTTGCAAAATTAATTAATGAAGTCTGTTTTTGAAGAACCTCGGTGGAGGA 121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
:||:|||||:::|||||
Db 122 AGTTTGGGAGTGTTCGAGCCAAATGGATATCACAGGACAAGGAGGTGGCTGTTAAG 181

QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
:||:|||||:::|||||
Db 182 AAGCTCCTCAAAGATAGAAAAGAGCGAGA-----ATA 214

QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrlGlyAlaCysLeuAsnPro--- 99
|||:::|||||:::|||||
Db 215 CTGAGTGCCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTTGAACCTCCC 274

QY 100 ---ValCysLeuValMetGluTyrlAlaGluGlyGlySerLeuTyrlAsnValLeuHisGly 118
:||:|||||:::|||||
Db 275 AACTATGTCATGTGCAGAAATATCTCTCTGGATCACTCTATGATTACATTAAACAGT 334

QY 119 -----AlaGluProLeuProTyrlThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
:||:|||||:::|||||
Db 335 AACAGAAATGAGGAGATG-----CATATGCATCACATTATGACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrlLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
:||:|||||:::|||||
Db 386 GTAGCCAAAGAAATGCAATATTATCATATGAGGCTCTCTGCAAGGCTATTTCACAGAGAC 445

QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
|||:::|||||:::|||||
Db 446 CTCAGTCAAGAAACGGTGTATAGCTGCTGATGA---GTACTGAAGATCTGTGACTTT 502

QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
|||:::|||||:::|||||
Db 503 GGT---GCCTCTCGGTTCATTAACATACACACACATGCTCCTTGGTTGGAACCTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrlSerGluLysCysAspValPheSer 214
|||:::|||||:::|||||
Db 560 TGGATGGCTCCAGAAATATCCAGAGTCTCCCTGTGCAAAACTTGTGACATATTC 619

QY 215 TrpGlyIleIleLeuTyrlGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
:||:|||||:::|||||
Db 620 TATGTTGGTTCCTCGGAGATGCTAACAGGAGGTCCCTTTAAAGTTTGGAGGA 679

QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
:||:|||||:::|||||
Db 680 -----TTCAAGTAGCTTGGCTGTAGTGAAAAAACAAGAGATTAACCATTCCAAGC 733

QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
:||:|||||:::|||||
Db 734 AGTTGCCCAAGATTTTGTCTGAACTGTATCATAGTTGGAGTGTGGAGCTGATGCCAAGAAA 793

QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrlPheProGly 293
:||:|||||:::|||||
Db 794 CGGCATCATTTCAAGCAATCATTTCAATCTG----- 826

QY 294 AlaAspGluProLeuGlnTyrlProCysGlnTyrlSerAspGluGlyGlnSerAsnSerAla 313
|||:::|||||:::|||||
Db 827 -----GAGTCCATCTCAATGACACG 847

QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
|||:::|||||:::|||||
Db 848 -----AGCCTTCCTGCAAGTGAATCATCTCTTACACAAACAGGCGGAGTGG 895

QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
|||:::|||||:::|||||
Db 896 AGGTGCGAA--ATTGAGGCAACTCTTGAGAGGCTAAAGAAACTAGAGCGGTGATCTCAGC 952

QY 354 LysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
:||:|||||:::|||||
Db 953 TTTAAGGAGCAGAGAGCTTAAAGAACAGAGAAAGAGCTTTAAAGATGTGGGAGCAAAAGCTG 1012

QY 371 SerHisLysSerValGluSerLeuProProThrSer----- 383
:||:|||||:::|||||
Db 1013 ACAGAGCAGTCCAACACCCCGCTTCTTGCCCTCTTGCTGCAAGAAATGTCTGAGGAGTCT 1072

QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAa 401
:||:|||||:::|||||
Db 1073 TACTTTGAATCTTAAACACAGAGGAGTCAAAACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132

QY 402 ThrThr----GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
:||:|||||:::|||||
Db 1133 ACAAGTAACGGGAGGCGCATGCATGAACCCAAAGTCTGCAGGCCATGATGCTGATGGGC 1192

QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerVal 435
|||:::|||||:::|||||
Db 1193 TTT-----GGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

RESULT 9
US-09-221-416-4
; Sequence 4, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-416-4

Alignment Scores:
Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: 3 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-416-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
|||:::|||||:::|||||
Db 2 TCGACCCACGGCTCGGTGGAGTATATACTTTGTTCATTATGAGATGTCTCTCGGT 61

QY 26 LeuAsnPheGluIleAspTyrlLysGluIleGluValValGlyArgGly 45
|||:::|||||:::|||||
Db 62 GCCTCCTTTTGCAAAATTAATTTGATGACTTCGCAATCTTTTGAAGAACCTCGGTGGAGGA 121

QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
:||:|||||:::|||||
Db 122 AGTTTGGGAGTGTTCGAGCCAAATGGATATCACAGGACAAGGAGGTGGCTGTTAAG 181

QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
:||:|||||:::|||||
Db 182 AAGCTCCTCAAAGATAGAAAAGAGCGAGA-----ATA 214

QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrlGlyAlaCysLeuAsnPro--- 99
|||:::|||||:::|||||
Db 215 CTGAGTGCCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTTGAACCTCCC 274

QY 100 ---ValCysLeuValMetGluTyrlAlaGluGlyGlySerLeuTyrlAsnValLeuHisGly 118
:||:|||||:::|||||
Db 275 AACTATGTCATGTGCAGAAATATCTCTCTGGATCACTCTATGATTACATTAAACAGT 334

QY 119 -----AlaGluProLeuProTyrlThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
:||:|||||:::|||||
Db 335 AACAGAAATGAGGAGATG-----CATATGCATCACATTATGACCTGGGCCACTGAT 385

QY 137 CysSerGlnGlyValAlaTyrlLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
:||:|||||:::|||||
Db 386 GTAGCCAAAGAAATGCAATATTATCATATGAGGCTCTCTGCAAGGCTATTTCACAGAGAC 445

QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
|||:::|||||:::|||||
Db 446 CTCAGTCAAGAAACGGTGTATAGCTGCTGATGA---GTACTGAAGATCTGTGACTTT 502

QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
|||:::|||||:::|||||
Db 503 GGT---GCCTCTCGGTTCATTAACATACACACACATGCTCCTTGGTTGGAACCTTCCCA 559

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrlSerGluLysCysAspValPheSer 214
|||:::|||||:::|||||
Db 560 TGGATGGCTCCAGAAATATCCAGAGTCTCCCTGTGCAAAACTTGTGACATATTC 619

QY 215 TrpGlyIleIleLeuTyrlGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
:||:|||||:::|||||
Db 620 TATGTTGGTTCCTCGGAGATGCTAACAGGAGGTCCCTTTAAAGTTTGGAGGA 679

QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
:||:|||||:::|||||
Db 680 -----TTCAAGTAGCTTGGCTGTAGTGAAAAAACAAGAGATTAACCATTCCAAGC 733

QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
:||:|||||:::|||||
Db 734 AGTTGCCCAAGATTTTGTCTGAACTGTATCATAGTTGGAGTGTGGAGCTGATGCCAAGAAA 793

QY 274 ArgProSerMetGlu


```

Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGTTCCATAAACACACACACATGCTTGGTTGGAACCTTCCCA 559

Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCAGAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGTGACACATATCC 619

Qy 215 TrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGGTGTGGTTCTCTGGAGATGCTAAACAAGGAGGCTCCCTTTAAAGGTTTGGAGGA 679

Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTATGTAAGAAACGAGAGATTAACCATTCGAAGC 733

Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTGTGCTGAACCTGTTACATCAGTGTTGGGAAGCTGATGCCAAGAA 793

Qy 274 ArgProSerMetGluGluIleValIleValIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTG-----826

Qy 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAAAATGACACG 847

Qy 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTCTCTGACAGTGAAGTAACTCACTCTCACAAACGCGGAGTGG 895

Qy 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeu 353
Db 896 AGGTGCGAA---ATTGAGGCAACTTGTAGAGGCTAAAGAACTAGACGCTGATCTCAGC 952

Qy 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGACGAGAGCTTAAAGAACGAGAAAGACGTTTAAAGATGTGGGCAAAAGCTG 1012

Qy 371 SerHisGlySerSerValGluSerLeuProProThrSer-----383
Db 1013 ACAGACAGTCCACACCCCGCTTCTTCTGCTTCTGCTGCAAGAAATGCTGAGGAGTCT 1072

Qy 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaAlaAla 401
Db 1073 TACTTTGAATCTAAACACAGAGGAGTCAACACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132

Qy 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAACGGGAGGCGCATGGCATGAACCCCAAGTCTCGAGGCCATGATGCTGATGGGC 1192

Qy 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

```

RESULT 11

```

US-09-163-115-4
; Sequence 4, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-163-115-4

Alignment Scores:
Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: 4 Gaps: 19

US-09-830-144-2 (1-579) x US-09-163-115-4 (1-2120)

Qy 10 SerSerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACCGGTCGGTGGAGATATAATCTTTGTCTATTATGAGATGTCCTCTCTCGGT 61

Qy 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluValValGlyArgGly 45
Db 62 GCCTCTTTGTGCAAAATTAATTTGATGACTTGCAGTTTGTGAAACTGCGGTGGAGGA 121

Qy 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGGAGTGTATTATCGAGCCAAATGGATATCACAGGACAAAGGAGGTGGCTGTAAG 181

Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGTCTCTCAAAATAGAGAAAGAGGCAGAA-----ATA 214

Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTTGAACCTCCC 274

Qy 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGTCATGTGCAGAAATATGCTTCTCTGGCATCACTCTATGATTACATTAAACAGT 334

Qy 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACGAGAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAAGTCATTATTATACATATGGAGCTCTCTCAAGTGTATTCACAGAC 445

Qy 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502

Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATAACCATACACACACATGCTTGGTTGGAACCTTCCCA 559

Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGTGACACATATTC 619

Qy 215 TrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGGTGTGGTTCTCTGGAGATGCTAAACAAGGAGGAGTCCCTTTAAAGGTTTGGAGGA 679

Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTATGTAAGAAACGAGAGATTAACCATTCGAAGC 733

Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGTTTGTGCTGAACCTGTTACATCAGTGTTGGGAAGCTGATGCCAAGAA 793

Qy 274 ArgProSerMetGluGluIleValIleValIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTG-----826

```

QY 294 AlaSerGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
DB 827 -----GAGTCATGTCAAATACACG 847
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
DB 848 -----AGCCTTCCTGCACAAAGTGTAACTATTCCTACACAAAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLeu 353
DB 896 AGGTGGGAA---ATTGAGCAACTCTTGGAGAGCTAAAGAACTAGACGTGATCTCAGC 952
QY 354 LysAsnGlnAlaLysGlnGlnSerGlnSerGlyArgLeuSerLeu-----GlyAla 370
DB 953 TTTAAGGACGACGAGCTTAAAGACGAAAGACCTTTAAAGATGTGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGlnSerLeuProProThrSer----- 383
DB 1013 ACAGAGCAATCCAAACACCCCGCTTCTTGCTTGGCAAGAATGCTGAGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 401
DB 1073 TACTTTGAATCTAAACACAGAGAGCAAGTCAAGATGTCATGTGATCAGACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
DB 1133 ACAAGTACGGGAGGCGCATGCGATGAAACCAAGTCTCAGCGCATGCTGATGGGC 1192
QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
DB 1193 TTT-----GGGATATCTTCTTCATGAACAAAGCAGAGCTGTG 1231

RESULT 12
US-09-221-528-4
Sequence 4, Application US/09221528
Patent No. 6190874
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MN1-050
CURRENT APPLICATION NUMBER: US/09/221,528
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1411)
US-09-221-528-4

Alignment Scores:

Pred. No.: 4, 97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.798 Conservative: 84
Best Local Similarity: 30.338 Mismatches: 159
Query Match: 16.264 Indels: 74
DB: 4 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-528-4 (1-2120)

QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCACCGCTCCGGTGAAGATATATCTTGTTCATTATGAGTGTCTCTCCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValGlyArgGly 45
DB 62 GCCTCCTTGTGCATAATTAATTGATGACTGAGTTTGTAAAGATGCGGAGTGAAGA 121

QY 46 AlaPheGlyValAlaCysLysAlaLysTyr-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGTGTATTATCGAGCCAAATGATATATACAGACAAAGGAGGTGGCTGAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AACCTCTCAAAATTAAGAAAGAGCGAGAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAAGTCTCTCACTACAGAAACATCATCCAGCTTTATGAGGTATTTGAACTCC 274
QY 100 ---ValCysLeuValMetGlyTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
DB 275 AACATGCACTTGCACAGAAATATGCTTCTCTGGATCACTATGATTAATTAACAT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerThrCysLeuGln 136
DB 335 AACAGAGTGAAGAGATG-----GATATGATCACTTATGACCTGGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLysLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTAGCCAAAGGATGCATTTTATCATATGAGAGCTCCTGTCAAGTATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
DB 446 CTCAAATCAAGAAACGTTTATATGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
DB 503 GGT---GCCCTCGGTTCCATTAACCAATACACACACATGCTTGGTGAACCTTTCCCA 559
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGTATGCTCCAGAAAGTTTATCCAGAGCTCCCGTGCACAAACCTGTGACATATTC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 234
DB 620 TATGTGTGGTCTCTGTGGAGATGCTTAACAAAGGAGGTCCCTTTAAAGGTGTGAAGA 679
QY 235 ProIaPheArgIleMetTyr---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAAGTACCTTGCTGTATGAGAAAAAACAAGAGATTACCATTTCAAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTyrPheLysAspProSerGln 273
DB 734 AGTTGCCCCAGAAAGTTTGTGGAAGCTGTATCATCAGTGTGGAGCTATGCCAAGAAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
DB 794 CGGCCATCATTCACAGCAAACTATTCATCTCTG----- 826
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlnSerAsnSerAla 313
DB 827 -----GAGTCCATGTCAAATGACAG 847
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
DB 848 -----AGCCTTCCTGCACAAAGTGTAACTATTCCTACACAAAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLeu 353
DB 896 AGGTGGGAA---ATTGAGCAACTCTTGGAGAGCTAAAGAACTAGACGTGATCTCAGC 952
QY 354 LysAsnGlnAlaLysGlnGlnSerGlnSerGlyArgLeuSerLeu-----GlyAla 370
DB 953 TTTAAGGACGACGAGCTTAAAGACGAAAGACCTTTAAAGATGTGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGlnSerLeuProProThrSer----- 383
DB 1013 ACAGAGCAATCCAAACACCCCGCTTCTTGCTTGGCAAGAATGCTGAGAGTCT 1072

```

Qy 384 -----GluGlyLysArgMetSerAlaAspMetSerGluLeuGluAlaAlaAla 401
Db 1073 TACTTTGAATCTAAACACAGAGAGTCAACAGAGTCGAGATGTCATGTCAGATCACAGCA 1132
Qy 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAAAGGGAGGCCATGGCATGAACCAAGTCTGCAGGCCATGATGCTGATGGGC 1192
Qy 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATCAACAAAGCAGGAGCTGTG 1231

RESULT 13
US-09-593-553-4
; Sequence 4, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-593-553-4

Alignment Scores:
Pred. No.: 4,97e-36 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: Gaps: 19

US-09-830-144-2 (1-579) x US-09-593-553-4 (1-2120)
Qy 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCGACCCACGCGTCCGGTGAAGTATAAATTTGTCATTATGAGATGTCGCTCTCGGT 61
Qy 26 LeuAsnGluGluLeuAspTyrLysGluLeuGluValGluValGlyArgGly 45
Db 62 GCCTCCTTTGTGCAAAATTAATTTGATGACTTGCAGTCTTTTGAACCTGCGGTGGAGGA 121
Qy 46 AlaPheGlyValValCysLysAlaLysTyr-----ArgAlaLysAspValAlaLys 63
Db 122 AGTTTGGGAGTGTATTCGAGCAATATGATATCATCAGCAAGAGGAGTGTCTGAAG 181
Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAAGAGGAGCA-----ATA 214
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTGAACCTCCC 274
Qy 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATGTCTCAGAAATATGCTTCTGGGATCACTCTATGATTACATTAAACAGT 334
Qy 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGln 136
Db 335 AACAGAGTGGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385

```

```

Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCAAAAGGAATCATATTTATCATATGAGGCTCTCTGCAAGGTGATTCACAGAGAC 445
Qy 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAAAGCTTGTATAGTCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCTCAATCAACACACATGCTCTTGGTTGGAACCTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluLysSerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAAAGTTATCCAGAGTCTCCTGTGTGAGAAACTTGTGACACATATTC 619
Qy 215 TrpGlyIleLeuLeuTyrGluValIleThrArgLysPropheAspGluIleGly 234
Db 620 TATGTTGTGTTCTCTGGGAGATGCTAAAGGAGGAGTCCCTTTAAAGGTTTGAAGGA 679
Qy 235 ProAlaPheArgIleMetTyr---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAAGTAGCTTGCTGTAGTGAAGAAACGAGAGATTAAACCATCCAAGC 733
Qy 254 AsnLeuProLysProLysLeuMetThrArgCysTyrSerLysAspSerGln 273
Db 734 AGTTGCCCCAGAAAGTTTGTGCTGAAGTGTATCATCAGTGTGGGAGCTGATGCCAAGAA 793
Qy 274 ArgProSerMetGluGluValIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTAAGCAAAATCATTTCAATCTCTG----- 826
Qy 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCAAAATGACACG 847
Qy 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCCTTCTCGCAAGTGTAACTCATCTCTACACAACAGGCGGAGTGG 895
Qy 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
Db 896 AGGTGCGAAA---ATTGAGGCAACTCTTCAGAGGCTAAAGAAACTAGAGCGTGTATCTCAGC 952
Qy 354 LysAsnGlnAlaLysGlnSerGluSerCysArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGACGAGGAGCTTAAAGAACGAGAACGCTTTAAAGATGTGGGAGCAAAAGCTG 1012
Qy 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
Db 1013 ACAGAGCAGTCCAACACCCCGCTTCTCTTGCCTCTGCTGCAAGAATGCTGTGAGGAGTCT 1072
Qy 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaAlaAla 401
Db 1073 TACTTTGAATCTAAACACAGAGGAGTCAACAGTGCAGAGATGTCATGTCAGATCACAGCA 1132
Qy 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAAAGGGAGGCCATGGCATGAACCAAGTCTGCAGGCCATGATGCTGATGGGC 1192
Qy 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATCAACAAAGCAGGAGCTGTG 1231

RESULT 14
US-09-593-553-4
; Sequence 4, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237

```

CURRENT FILING DATE: 1998-12-28
 EARLIER APPLICATION NUMBER: 09/163,115
 EARLIER FILING DATE: 1998-09-29
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 4
 LENGTH: 2120
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (47) .. (1411)
 US-09-221-237-4

Alignment Scores:
 Pred. No.: 4.97e-36 Length: 2120
 Score: 490.00 Matches: 138
 Percent Similarity: 48.79% Conservative: 84
 Best Local Similarity: 30.33% Mismatches: 159
 Query Match: 16.26% Indels: 74
 DB: 4 Gaps: 19

US-09-830-144-2 (1-579) x US-09-221-237-4 (1-2120)

10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
 2 TCGACCCAGCGCGCGCGGAGATTAATACCTTGTCAATATGAGATGCGCTCGGT 61
 26 LeuAsnPhcGluGluIleAspTyrIysGluIleGluValGluValValGlyArgGly 45
 62 GCCTCCTTGTGCAAAATTAAATTGATGACTTGCACTTTTGGAAACGCGGTGAGGA 121
 46 AlaPhcGluValValCysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
 122 AGTTTGGAGGCTTATATGAGCCAAATGATATCAGACAGACAGAGGTGCTGTAAG 181
 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
 182 AAGCTCCTCAAAATAGAGAAAGAGCAGAA-----ATA 214
 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
 215 CTCAGAGTCTCTCAGCAGAAACATCATCATGTTTATGAGGATATTTGAACTCTCC 274
 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
 275 AACTATGCAATTCACAGAAATATGCTTCTGCGGATCATCTATGATATACATTAAAGT 334
 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
 335 AACAGAAAGTGAGGAGATG-----GATATGATCATCATTAATGACCTGGGCCACTGAT 385
 137 CysSerGlnGlyAlaAlaTyrIleuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
 386 GTAGCCAAAGGAATGCAATTATTAATGATGAGGCTCTGTCAGAGGTATCAACAGAGC 445
 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
 446 CTCAGAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAGATCTGTGCTTT 502
 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
 503 GGT---GCTCTCGGTTCCATACACATACACACATGCTCTTGAGTGAACCTTCCCA 559
 195 TyrMetAlaProGluValAlaPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
 560 TGGATGGCTCCAGAAAGTTATCCAGAGTCTCTGCTGTCGAAACCTTGTAACATATTC 619
 215 TyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
 620 TATGGTGTGTTCTCTGAGAGATGCTAACAGAGGAGGTCCCTTTAAAGTTTGAAGA 679
 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253

680 -----TTCAAGTACTGCTGCTTGTAGTGAAAAAAGAGAGTTTAACTTCAAAC 733
 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
 734 AGTTGCCCAAGAGTTTCTGTAACGTATCATCATGTTGAGGAGCGATCCCAAGAA 793
 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
 794 CGGCCATCATTCACAGCAAAATCATTTCAATCCCTG----- 826
 294 AlaAspGluProLeuIleTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
 827 -----GAGTCCATGTCATTAATACACG 847
 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
 848 -----AGCTTCTCTGACAAAGTGAATCTATCTTCTCAACAAAGCGGAGTGG 895
 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
 896 ACGTGCAGAA---ATTGAGGCAACTCTTGAGAGGCTTAAAGAACTAGAGCGTATCTCAGC 952
 354 LysAsnGlnAlaLysGlnGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
 953 TTTAAGAGCAGAGGCTTAAAGAACAGAGAAAGAGCTTTAAAGATGTGGAGCAAAAGCTG 1012
 371 SerHisGlySerSerValGluSerLeuProProThrSer----- 383
 1013 ACAGAGAGCTCAACACCCCGCTTCTTGCTGCTGCTGCAAGAAATGTCTGAGAGTCT 1072
 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla 401
 1073 TACTTGAATCTTAAACAGAGAGATCAACAGTCAAGATGATGATGATGATGATGATGATGAT 1132
 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
 1133 ACAAGTACCGGAGGCGCCATGAGCATGAAACCAAGTCTGAGCCCATGATGCTGATGGGC 1192
 421 ThrGluProGluGlnValSerSerArgSerSerSerProSerVal 435
 1193 TTT-----GGGATATCTTCTCAATGAAACAAAGCAGAGCTGTG 1231

RESULT 15
 US-09-221-235-6
 Sequence 6, Application US/09221235
 Patent No. 6043040
 GENERAL INFORMATION:
 APPLICANT: Acton, Susan
 TITLE OR INVENTION: NOVEL CSARK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
 FILE REFERENCE: NMT-050
 CURRENT APPLICATION NUMBER: US/09/221,235
 EARLIER FILING DATE: 1998-12-28
 EARLIER APPLICATION NUMBER: 09/163,115
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 6
 LENGTH: 1365
 TYPE: DNA
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1) .. (1365)
 US-09-221-235-6

Alignment Scores:
 Pred. No.: 1.69e-35 Length: 1365
 Score: 481.00 Matches: 132
 Percent Similarity: 49.08% Conservative: 81
 Best Local Similarity: 30.41% Mismatches: 151
 Query Match: 15.96% Indels: 70
 DB: 3 Gaps: 18

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 23:04:29 ; Search time 83.4017 Seconds
(without alignments)
2707.033 Million cell updates/sec

Title: US-09-830-144-2

Perfect score: 3014

Sequence: 1 MSTAASASSSSSSAGEMIE.....QCKKOLEVIRSQQKROGTS 579

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141354_2307/app_query.fasta_1.1422
-DB=Published_Applications_NA -QFWT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09830144 @CGN 1 1 51 @runat_04122002_141354_2307
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DSV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3014	100.0	1788	9	US-10-158-895-14
2	3014	100.0	2656	9	US-10-158-895-3
3	493.5	16.4	3454	10	US-09-969-347-226
4	490	16.3	2120	10	US-09-757-982-4

ALIGNMENTS

RESULT 1

US-10-158-895-14
; Sequence 14, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-10-158-895-14

Sequence 6, Appli
Sequence 1, Appli
Sequence 3, Appli
Sequence 1, Appli
Sequence 7, Appli
Sequence 9, Appli
Sequence 1014, Ap
Sequence 1, Appli
Sequence 1577, Ap
Sequence 1073, Ap
Sequence 882, App
Sequence 16, Appl
Sequence 903, App
Sequence 84, Appl
Sequence 405, App
Sequence 2, Appli
Sequence 48, Appl
Sequence 271, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 1085, Ap
Sequence 1, Appli
Sequence 1405, Ap
Sequence 491, App
Sequence 4, Appli
Sequence 23, Appl
Sequence 1, Appli
Sequence 207, App
Sequence 1, Appli
Sequence 3, Appli
Sequence 553, App
Sequence 1, Appli
Sequence 1479, Ap
Sequence 1128, Ap
Sequence 34, Appl
Sequence 7, Appli
Sequence 13, Appli

Alignment Scores:

Pred. No.:	5,93e-251	Length:	1788
Score:	3014.00	Matches:	579
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2 (1-579) x US-10-158-895-14 (1-1788)

```

QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20
DB 7 ATGCTCAAGCCCTCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 66
QY 21 AlaProSerGlnValLeuAsnPhelGluIleAspTyrIleGluValGluGlu 40
DB 67 GCGCCCTTCCAGAGTCTCTCAACTTGAAGAGATGACTACAGAGAGATCGAGGTGGAAG 126
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrAlaLysAspVal 60
DB 127 GTTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 186
QY 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 187 GCTATTAAACAAATAGAAAGTGAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAG 246
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyValaCysLeuAsnProVal 100
DB 247 TTATCCCGTGTGAACCATCTTAATATTGTAAACCTTTATGAGAGCCGCTTGAAATCCAGTG 306
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
DB 307 TGTCTGTGATGAAATATGCTGAAGGGGGGCTTTTATATATGCTGATGCTGCTGAA 366
QY 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
DB 367 CCATTCGCATATTATCTGCTGCCACGCATAGTGGTGTGTACAGTGTCTCCAGAGA 426
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 427 GTGGCTTATCTTACAGCAAGCAACCAAGCCCTAATTCACAGGAGACCTGAAACCAACA 486
QY 161 AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB 487 AACTTACTGCTGCTGAGGGGGGAGACAGTCTTAAAAATTTGTGATTTTGTGACAGCTGT 546
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
DB 547 GACATTCAGACACACATGACCAATTAACAGGGAGTGTGCTGTGATGGACCTGAAAGTT 606
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyr 220
DB 607 TTGGAAGGTGATATTACATGAAAAATGTGAGCTTCAAGCTGGGGATATTATCTTTGG 666
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
DB 667 GAAGTGATAACGGCGTCGGAACCCCTTTGATGAGATGTGGGCCCACTTCCCAATCAAG 726
QY 241 TyrAlaAlaHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 727 TGGGCTGTATATATGATGACTCGAACCACTATATAAAAAATTTACTTAAGCCCATTTGAG 786
QY 261 SerLeuMetThrArgCysTyrPserLysAspProSerGlnArgProSerMetGluGluIle 280
DB 787 AGCCTGATGACTGCTTGTGTCTAAAGATCCTTCCAGGCGCCCTTCAATGAGGAAATTT 846
QY 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyValaAspGluProLeuGlnTyr 300
DB 847 GTGAATAATATGACTACTGATGCGGTACTTCCAGGAGCAAGATGAGCATTAACAGTAT 906
QY 301 ProCysGlnTyrSerArgGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet 320

```

```

DB 907 CCTGTGATGATTTCAGATGAGAGACAGAGCACTCTGCCACGATCAGGCTCATTCATG 966
QY 321 AspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGluValProAla 340
DB 967 GACATTCCTCTCAACAAATATGACATCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1026
QY 341 ThrAsnAspThrIleLysAsnGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln 360
DB 1027 ACAATATATCTATTAAAGCCCTTAGAATCAAAATGTGTGAAAAATCAAGCAAGCAAGCAAG 1086
QY 361 SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro 380
DB 1087 AGTGAATCTGAGCCTTTAACTTGGAGAGCTCCCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1146
QY 381 ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla 400
DB 1147 CCAACTCTGAGGCGCAAGAGAGATGAGTGCATGCAATGCTGMAATAGAAAGTACAGATCGCC 1206
QY 401 AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
DB 1207 GCAACCAAGGCAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1266
QY 421 ThrGluProGlyGlnValSerSerArgSerSerSerProSerValArgMetIleThrThr 440
DB 1267 ACAGAACTGCTGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1326
QY 441 SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr 460
DB 1327 TCAGAGACCAACTCAGAAAGCCAACTCGAAATCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1386
QY 461 AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu 480
DB 1387 GATACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
QY 481 GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys 500
DB 1447 CAGCCTCTAGACAGCGTCCCAACTCCAAAGAACTATATGACAGTGTGAAACGACATGT 1506
QY 501 LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys 520
DB 1507 AAAATGGCACAAGATATATGAAAGTTCAAAACAGAAATTCATGTTATATACAGAGAAAG 1566
QY 521 GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu 540
DB 1567 CAAGAACTAGTTCAGAACTGAGACAGAGATGAAAAAGACAGCAAAATACATCTCGCTG 1626
QY 541 ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln 560
DB 1627 GTACAGGAACATTAAGCTTTTATGATGAAAAACAAAGCCCTTCTACTTACACAGCAA 1686
QY 561 CysLysLysGlnLeuGluValIleArgSerGlnGlnGlnLysArgGlnGlnGlnGlnGln 579
DB 1687 TGCAAAAAACAATGAGAGTGCATCAGAAATCAGACAGCAAGCAAGCAAGCAAGCAAGCAAG 1743

```

RESULT 2

```

US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48

```



```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-158-895-3

```

Alignment Scores:

Pred. No.:	1.02e-250	Length:	2656
Score:	3014.00	Matches:	579
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2 (1-579) x US-10-158-895-3 (1-2656)

QY	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu	20
Db	183	ATGTTCTACAGCCTCTGCGCGCTCTCTCTCTCTCTCTCTCTCGCCGCTGAGATCGAA	242
QY	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrIysGluValGluGlu	40
Db	243	GCCCTCTCCAGGTCCTCAACTTTGAAGAGATCGACTACAAAGAGATCGAGGTGGAAGAG	302
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaIysTrpArgAlaLysAspVal	60
Db	303	GTTGTTGGAAGAGAGAGCCTTTGGAGTTGTTTGCAAAGCTAAAGTCGAGAGCAAAAGATGTT	362
QY	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	363	GCTATTAAACAAATAGAAAGTGAAATCTGAGAGGAACGCTTTATTGTAGAGCTTCGGCAG	422
QY	81	LeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCysLeuAsnProVal	100
Db	423	TTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATCGAGCCTGCTTGAAATCCAGTG	482
QY	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	483	TGCTCTTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTCGTCATGCTGCTGAA	542
QY	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysIleuGlnCysSerGlnGly	140
Db	543	CCATTGGCCATTATATCTGCTGCCACGCAATGAGTTGGTGTGTTACAGTGTCCCAAGGA	602
QY	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160
Db	603	GTGGCTTATCTTCACGCATGCAACCCAAAGCGCTAATTCACAGGACCTGAAACACCA	662
QY	161	AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	663	AACTTACTGTGTTGAGGGGGGACAGTTCTAAAAATTTGCTGATTTGGTACAGCCTGT	722
QY	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	723	GACATTACAGACACATGACCAATAACAAGGGAGTGCTCTTGGATGGCACCTGAAGTT	782
QY	201	PheGluClySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp	220
Db	783	TTTGAGGGTAGTAATACAGTGAATAATGTGACGCTTCACGTGGGGCTATTATCTTTGG	842
QY	221	GluValIleThrArgArgLysProPheAspGluIleClyGlyProAlaPheArgIleMet	240
Db	843	GAAATGATAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCACGCTTCCGAATCATG	902
QY	241	TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	903	TGGCTGTTCAATATGTACTCGCAACCACTACTATAAAAAATTTACTAAGCCCATTTAG	962
QY	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArqProSerMetGluGluIle	280

Db	963	AGCCTGATGACTCGTTGTTGGTCTTAAGAGATCCTTCCACGCGCCCTTCAATGGAGAAATT	10322
Qy	281	VallysileMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	1023	GTGAAATAATGACTCACTTGATCGGTACTTTCAGGAGCAGATGAGCCATTACAGTAT	1082
Qy	301	ProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThrSerThrGlySerPheMet	320
Db	1083	CCTTGTCAGTATTCAAGATGAAGACAGACAGCAACTCTGCCACCAGTACAGGCTCATTCATG	1142
Qy	321	AspileAlaSerThrAsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAla	340
Db	1143	GACATTGCTTCAAAATACGATGAAACAAAGTAGCACTAATATCGAGCAAGTTCCTGCC	1202
Qy	341	ThrAsnAspThrIleLysArgLeuGluSerLysLeuLeuLysAsnGlnAlaLysGlnGln	360
Db	1203	ACAAATGATACTATTAAAGCGCTTAGAATCAAAATTTGTAATAATCAGGCAAGCAACAG	1262
Qy	361	SerGluSerGlyArgLeuSerLeuGlyAlaSerHisGlySerSerValGluSerLeuPro	380
Db	1263	AGTGAATCTCGAGCTTTAAAGTTGGAGAGCTCCCATGGGAGCAGTGTGGAGAGCTTGCCC	1322
Qy	381	ProThrSerGluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAla	400
Db	1323	CCAACTTCTAGGGCAAGAGGATGAGTGTCTGACATGTCTGAAATAGAAGCTTAGATCGCC	1382
Qy	401	AlaThrThrGlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly	420
Db	1383	GCAACCAAGCAACGACAGCCAGACGCTAGATCCATCCAAAGACTTCGACTGTAACGGGA	1442
Qy	421	ThrGluProGlyGlnValSerSerArgSerSerProSerValArgMetIleThrThr	440
Db	1443	ACAGAACCTGGTCAGGTGACAGTAGGTTCATCCAGTCCCGAGTGCAGAAATGATTACTACC	1502
Qy	441	SerGlyProThrSerGluLysProThrArgSerHisProTyrThrProAspAspSerThr	460
Db	1503	TCAGGACCACTCAGAAAAGCCAACTCGAAGTCATCCATGGACCCCTGATGATTCCACA	1562
Qy	461	AspThrAsnGlySerAspAsnSerIleProMetAlaTyrLeuThrLeuAspHisGlnLeu	480
Db	1563	GATACCAATGGATCAGATAACTCCATCCCAATGGGTATTCTTACACTGGATCAGCAACTA	1622
Qy	481	GlnProLeuAlaProCysProAsnSerLysGluSerMetAlaValPheGluGlnHisCys	500
Db	1623	CAGCCTCTACACCGTCCCAAACTCCAAAGAAATCTATGGCAGTGTTTGAACAGCATGT	1682
Qy	501	LysMetAlaGlnGluTyrMetLysValGlnThrGluIleAlaLeuLeuLeuGlnArgLys	520
Db	1683	AAAATGGCACAGAATATATGAAAGTTCAACAGAAAATTGCATTCTTTATACAGAGAAAG	1742
Qy	521	GlnGluLeuValAlaGluLeuAspGlnAspGluLysAspGlnGlnAsnThrSerArgLeu	540
Db	1743	CAGAACCTAGTTCGAACCTGGACACAGGATGAAAGAGACCCAGCAAAATACATCTCGCGCTG	1802
Qy	541	ValGlnGluHisLysLysLeuLeuAspGluAsnLysSerLeuSerThrTyrTyrGlnGln	560
Db	1803	GTACAGGAACATAAAAGCTTTTAGATGAAACAAAAGCCCTTCTACTTACTACAGCA	1862
Qy	561	CysLysLysGlnLeuGluValIleArgSerGlnGlnLysArgGlnGlyThrSer	579
Db	1863	TGCAAAAAACCACTAGAGGTCTACAGAGTCAGAGTCAGAGCAGAGAACCCAGGACGCTTCA	1919

RESULT 3

```

US-09-969-347-226
; Sequence 226, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969.347
;

```

```

CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/60/237,598
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: US/60/237,604
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 318
SOFTWARE: PatentIn version 3.0
SEQ ID NO 226
LENGTH: 3454
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-347-226

Alignment Scores:
Pred. No.: 2,75e-33 Length: 3454
Score: 493.50 Matches: 145
Percent Similarity: 46.49% Conservative: 87
Best Local Similarity: 29.06% Mismatches: 191
Query Match: 16.37% Indels: 79
DB: 10 Gaps: 14

US-09-830-144-2 (1-579) x US-09-969-347-226 (1-3454)
QY 21 AlaProserGluValLeuAsnPhgIuGluIleAspTyLysGluIleGluValGluIu 40
DB 535 GCACCCGCGGCTCCAGCTGCCCCAGAGATCCCTTCACAGCTGACGCTAGAGAG 594
QY 41 ValValGlyArgGlyAlaPhgIuValValCysLysAlaLysTrpArgAlaLysAspVal 60
DB 595 ATCATCGTGTGGGGGCTTTGGCAAGGTCTATCGGGCCCTGTGGCGGAGAGGTG 654
QY 61 AlilLeuGlnIleGluSerGluArgLysAlaPheIleVal----- 76
DB 655 GCAGTCAAGGCCCCCGGCTGAGACCTTGAGAGACCCGCGAGTGCACGAGAGAGTG 714
QY 77 -----GluLeuArgGluLeuSerArgValAsnHisProAsnIleValLysLeuTyrgly 94
DB 715 TGGCAGAGAGCCCGGCTCTTTGGAGCCCTGACAGCCCAACATATGCTCCCTTAGGGGC 774
QY 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyraLagIuLysSerLeu 112
DB 775 GCGTGGCTCAACCCCGCACACCTCTGCTAGTATGAGATGATCCCGGGGTGGTGCATG 834
QY 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrrTyrrAlaAlaHisAlaMetSer 132
DB 835 AGCAGGAGTGTGGCAGAGTGCACCGGTGCCACCTCACGTG-----CTGGTCAAC 882
QY 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrrLeuHisSerMetGlnProLysAlaLeu 152
DB 883 TGGGCTGTGAGGTGGCCCGGGGATGAACCTACTACACATGATGCCCTGTGCCCATC 942
QY 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
DB 943 ATCCACCGGAGCTCAAGTCCATCAACATCTGATCTCTGAGGCCATCGAGAACCAAC 1002
QY 166 AlaGlyCylThrValLeuLysIleCysAspPheGlyThrAlaCysAsp--IleGlnThr 184
DB 1003 CTCGCAGACACGCGTGTCAAGATCACGAGACTTGGCCTGCCCGCAGAGTGCACAGAAC 1062
QY 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProLysValPheGluGlySer 204
DB 1063 ACCAAGATGAGCGCTGGGAGCTTACGCTGTGATGGCGCGGAGGTATCCGCTCTTCC 1122
QY 205 AsnTyrrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
DB 1123 CTCTCTCCAAAGACATGATGCTGTGAGGCTTGGGGTGTCTGTGGAGCTGCTGAGAG 1182
QY 225 ArgArgLysProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaVal--- 243
DB 1183 GGGAGAGTCCCTTACCGTGAGATC-----GACGCTTGGCGTGGCGATAGCGGTGGCT 1236
QY 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263

```

```

DB 1237 ATGAATAAGCTGACGCTGCCCATTTCCCTCCACGTCGCCGAGCCCTTTGGCCGCTCTG 1296
QY 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluIleValLysIle 283
DB 1297 GAGGAATCTGGGAGCCAGCCAGCCCGGCGGCGAGATTTCCGTAGCATCTTGAACCG 1356
QY 284 MetThrHisLeuMetArgTyrrPhePro----- 292
DB 1357 CTTAAGTCAATCAACAGTC-AGCCTGTTCAGATGCCATGAGTCTTCACATCGCT 1415
QY 292 ----- 292
DB 1416 GCAGAAAGACTGGAAGCTGAGATTACAGACATGTTGATGACCTTCGACCAAGAGAA 1475
QY 293 GlyAlaAspGluProLeuGlnTyrrProCysGlnTyrrSerAspGluLysIleSerAsnSer 312
DB 1476 GGAAGCTTCGAGCGCTG-AGAGAGAGCTGTGGCGGCGGACAGAGACAGCGCTTCAGG 1534
QY 313 AlaThrSerThrGly-----SerPheMetAspIleAlaSerThrAsnThrSer 328
DB 1535 AGGAGCAGCTGGCGGCGGAGCAGAGCTGGCAGAACTGAGATGACATCTGAGAAC 1594
QY 329 AsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeu 348
DB 1595 GCGAGCTGCACCTGCTC-ATGTGCGAGCTGAGCCAGAGAGAGCCCGGCTCGCAAGCGC 1653
QY 349 GluSerLysLeuLysAsnGlnAlaLysGlnGlnSerGluSerGly-----Arg 365
DB 1654 AAGGCAACTTCAAGCCACGCGCTGCTCAAGCTGCGGAGAGCGGCGGACCATCAGC 1713
QY 366 LeuSerLeuGlnAlaSerHisGlySerSerValGluSerLeuProProThrSerGluGly 385
DB 1714 CTGCCCTCTGGCTTTGACGCAATGATCAAGTCAAGCTCTTCAACTCTGAT----- 1767
QY 386 LysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaIleThrThrGlyAsn 405
DB 1768 AAGCGAAGAGATCGATGGGCGCAGCCCTGCAAGCCCGCATCATC----- 1818
QY 406 GlyGlnProArgArgArgSerIleGlnAspLeuThrValThrGlyThrGluProGlyGln 425
DB 1819 -----CCCGCGCTGAGAGGCGCATTCG-----CTGACTCCGCGTGTGGTGGC 1863
QY 426 ValSerSerArgSerSerProSerValArgMetIleThrThrSerGlyProThrSer 445
DB 1864 AGCAGCAGTGCAGCAGCAGAGAGAGAGAGAGATGGAGCGCGGTGGGCCCAAG 1923
QY 446 GluLys-----ProThrArgSerHisProTrpThrProAspAspSer 459
DB 1924 AAGGAAGAATGTCGTGGGGGCGCAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1980

RESULT 4
US-09-757-982-4
Sequence 4, Application US/09757982
Patent No. US20020094559A1
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/757,982
CURRENT FILING DATE: 2001-01-10
PRIOR APPLICATION NUMBER: 09/163,115
PRIOR FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 2120
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (47)..(1411)
US-09-757-982-4

```

Alignment Scores:

```
Pred. No.: 2,84e-33 Length: 2120
Score: 490.00 Matches: 138
Percent Similarity: 48.79% Conservative: 84
Best Local Similarity: 30.33% Mismatches: 159
Query Match: 16.26% Indels: 74
DB: 10 Gaps: 19

US-09-830-144-2 (1-579) x US-09-757-982-4 (1-2120)

QY 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db 2 TCAGCCACCGCTCGCGTGAAGTAAATCTTTGTCAATTATGATGATGCTCTCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluValGlyArgGly 45
Db 62 GCCTCCTTTGTCAATTAATTAATTTGATGACTTGCAGTTTGTGAAACATCGCGTGAGGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db 122 AGTTTGGGAGTGTATTATCAGGCCAATGATATCACAGGACAAGGAGGTGGCTGTAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCAAAATAGAGAAGAGGCAGAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTAGTGTCTCAGTCACAAACATCATCCAGTTTATGGAGTAATTTGTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTGTTCAGAGAATATGCTTCTCTGGATCACTCTATGATTACATTACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAAAGTGAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGAAATGCATTATTACATATATGAGGCTCCTCTCAAGGTGATTTCAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATAACCAACATACACACATGTCCTTGTGTGAACTTTCCCA 559
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGCTCCAGAAGTATCCAGAGTCTCCGTGTGTCAGAACTTGTGACATATATCC 619
QY 215 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 620 TATGTTGGTGTCTCTGGAGATGCTCAACAGGAGGTCCCTTTAAAGTTTGGGAAGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
Db 680 -----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACGAGAGATTAAACCATTCACAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTTGCCCCAGAGATTGTCGAACTGTTTACATCAGTGTGGGAAGCTGATGCCAAGAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly 293
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTCG----- 826
QY 294 AlaAspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAla 313
Db 827 -----GAGTCCATGTCATAATGACACG 847
```

```
QY 314 ThrSerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThr 333
Db 848 -----AGCTTCTCTGACAAAGTGAACCTCATCTCTACACAACAGCGGAGTGG 895
QY 334 AsnMetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGluSerLysLeuLeu 353
Db 896 AGGTGCGGAA--ATTGAGGCAACTCTTTGAGAGGCTTAAAGAACTAGAGCGTGATCTCAGC 952
QY 354 LysAsnGlnAlaLysGlnSerGluSerGlyArgLeuSerLeu-----GlyAla 370
Db 953 TTTAAGGACGAGGAGCTTTAAAGAACGAGAACGCTTTAAAGATGTGGGAGCAAAAGCTG 1012
QY 371 SerHisGlySerSerValGluSerLeuProThrSer----- 383
Db 1013 ACAGAGCAGTCCAACACCCCGCTTCTCTTGCTCTTGCTGCAAGAAATGCTGAGGAGTCT 1072
QY 384 -----GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAla 401
Db 1073 TACTTTGAATCTAAACACAGAGGAGTCAAAACAGTGCAGAGATGTCTATGTCAGATCACAGCA 1132
QY 402 ThrThr---GlyAsnGlyGlnProArgArgSerIleGlnAspLeuThrValThrGly 420
Db 1133 ACAAGTAAACGGGAGGCGCCATGGCATGAACCCAAAGTCTGCAGGCCATGATGCTGATGGC 1192
QY 421 ThrGluProGlyGlnValSerSerArgSerSerProSerVal 435
Db 1193 TTT-----GGGGATATCTTCTCAATGAACAAAGCAGGAGCTGTG 1231

RESULT 5
US-09-757-982-6
; Sequence 6, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-757-982-6

Alignment Scores:
Pred. No.: 9.32e-33 Length: 1365
Score: 481.00 Matches: 132
Percent Similarity: 49.08% Conservative: 81
Best Local Similarity: 30.41% Mismatches: 151
Query Match: 15.96% Indels: 70
DB: 10 Gaps: 18

US-09-830-144-2 (1-579) x US-09-757-982-6 (1-1365)

QY 27 AsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGlyAla 46
Db 19 TCCTTTGTGCAAAATTAATTTGATGACTTGCAGTTTTTTTGAAAACTGCGGTGGAGGAAGT 78
QY 47 PheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys--- 63
Db 79 TTTGGAGTGTATTTCGAGCCAAATGGATATCACAGGACAAGGAGGTGGCTGTAAAGAG 138
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeu 81
Db 139 CTCCTCAAAATAGAGAAAGAGGCAGAA-----ATATCTC 171
```

QY 82 SerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro----- 99
 DB 172 AGTGTCTCTCAGTACAGAAACATCATCTCCAGTTTATGGATTAATTTTAACTCCCAAC 231
 QY 100 ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly--- 118
 DB 232 TATGGCATTTCTCAGAAATATGCTTCTCTGGGATCTCATCTATGATTAACATTAAC 291
 QY 119 ---AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
 DB 292 AGAAGTGAAGAGATG-----GATATGGATCATATTATGACCTGGGCACTGATGTA 342
 QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
 DB 343 GCCAAAGGAATGATTAATTATCATATGAGAGCTCTCTGCAAGGTGATTCACAGAGCCCTC 402
 QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
 DB 403 AAGTCAAGAAACCTGTTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTTGGT 459
 QY 178 ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAlaTrp 195
 DB 460 ---GCCCTCGGTTCATACCATTAACAACACATGTCCTGGTGGAATTTCCCATGG 516
 QY 196 MetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTrp 215
 DB 517 ATGGCTTCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAAACCTGTGACACATATTCCTAT 576
 QY 216 GlyIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
 DB 577 GGTGTGTTCTCTGGGAGATGCTTAACAAGGAGAGTCCCTTTAAAGCTTTGGAAGA--- 633
 QY 236 AlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
 DB 634 ---TTACAATAGCTTGCTGCTAGTGAAGAAACAGAGAGATTAACCATTCACACAGCT 690
 QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
 DB 691 TGCCCCAGAAAGTTTGTGAACTGTTTACATCAGTGTGGAGAACCTGATCCAAAGAACCG 750
 QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgLysrPheProGlyAla 294
 DB 751 CCATCATTCAGCAAAATCATTTCAATCTG----- 780
 QY 295 AspGluProLeuGlnTyrProCysGlnTyrSerAspGluGlyGlnSerAsnSerAlaThr 314
 DB 781 -----GAGTCCCATGCAATGACAGC--- 801
 QY 315 SerThrGlySerPheMetAspIleAlaSerThrAsnThrSerAsnLysSerAspThrAsn 334
 DB 802 -----AGCCTTCTCTGCAAGTGAATCTCTTACACAAACAGCGGAGTGGAG 852
 QY 335 MetGluGlnValProAlaThrAsnAspThrIleLysArgLeuGlnSerLysLeuLys 354
 DB 853 TGGCAA---ATTAGGCAACTCTTGAGAGGCTTAAAGAAACTAGAGCGTGAATCTCAGCTTT 909
 QY 355 AsnGlnAlaLysGlnSerGlnSerGlyArgLeuSerLeu-----GlyAlaSer 371
 DB 910 AAGAGACAGAGACTTAAAGAACGAAAGACCTTTAAAGATGTGGAGACAAAGACTGACA 969
 QY 372 HisGlySerSerValGlnSerLeuProProThrSer----- 383
 DB 970 GAGAGTCCAAACCCCGCTTCTCTGCTCTTGCAAGATGTCTGAGAGTCTTAC 1029
 QY 384 ---GluGlyLysArgMetSerAlaAspMetSerGluIleGluAlaArgIleAlaAlaThr 402
 DB 1030 TTGGAATCTTAAACAGAGAGTCAACAGTGCAGAGATGTCATGATCAGACAGAAAC 1089
 QY 403 Thr---GlyAsnGlyGlnProArgArgArgSerIleGlnAspLeuThrAlaThrGlyThr 421
 DB 1090 AGTAAAGGAGGAGGCCATGCGATGAACCCCAAGTCTGAGGCCATGATGCTGATGGCTTT 1149

QY 422 GluProGlyGlnValSerSerArgSerSerSerProSerVal 435
 DB 1150 -----GGGATATATCTTCTCATATGAACAAAGACAGAGACTGTG 1185
 RESULT 6
 US-10-014-882-1
 ; Sequence 1: Application US/10014882
 ; Patent No. US20020107384A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Hu, Yi
 ; APPLICANT: Kieke, James
 ; APPLICANT: Donoho, Gregory
 ; TITLE OF INVENTION: No. US20020107384A1e1 Human Kinase and Polynucleotides Encoding r
 ; FILE REFERENCE: LEX-0279-USA
 ; CURRENT APPLICATION NUMBER: US/10/014,882
 ; PRIOR FILING DATE: 2001-12-11
 ; PRIOR FILING DATE: 2000-12-11
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3111
 ; TYPE: DNA
 ; ORGANISM: homo sapiens
 US-10-014-882-1
 Alignment Scores:
 Pred. No.: 1,37e-30 Length: 3111
 Score: 461.50 Matches: 176
 Percent Similarity: 41.16% Conservative: 94
 Best Local Similarity: 26.83% Mismatches: 238
 Query Match: 15.31% Indels: 149
 Gaps: 25
 US-09-830-144-2 (1-579) x US-10-014-882-1 (1-3111)
 QY 22 ProSerGlnValIleAsnPheGluGluIleAspTyrLysGluIleGluValIleGluVal 41
 DB 328 CCTCGCGGCCAGCTCCCGGTACAGTGCCTTCGAGGCGGTGAGCTGAAGAGACTC 387
 QY 42 ValGlyArgGlyAlaPheGlyValIleCysLysValAlaLysTrpArgAlaLysAspValAla 61
 DB 388 ATCGGCGGTGGGCTTCGGGAGGTGACCGGCCACTGGCAGGCGCCAGAGGTGGC 447
 QY 62 IleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
 DB 448 GTAAAGGCGCGCCAGACCCGAGACAGAGACGAGCGGCGGCTGCCAGAGGTGGC 507
 QY 77 ---GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95
 DB 508 CGCAGGCTCGCTCTTCCGATGCTGCGGACCCCAACATCATCGAGCTGCGGCGGTG 567
 QY 96 CysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyr 113
 DB 568 TGCCTGCAGCAGCCGACACTGTGCTGTGTGTGAGAGTTCCGCCGCGGAGGCTCAAC 627
 QY 114 AsnValLeuHisGlyAlaGluProLeuProTyrThrAla----- 127
 DB 628 CGAGCGCTGGCGGTGCCAACCGCGCCCGGACCCGCGCGCGCGCGCGCGCGCGCG 687
 QY 128 -----AlaHisAlaMet---SerTrpCysLeuGlnCysSerGlnGlyValAla 142
 DB 688 CGCGCATCCCTCCGACAGTGTGTCGACAGTGGCGCTGACAGATAGCGCGGCGCATCTC 747
 QY 143 TyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuProProAsnLeu 162
 DB 748 TACCTGATGAGAGGCTTCGTGCTCCATCTGACCGGAGACTCAAGTCCAGCAAACTT 807
 QY 163 LeuLeuValAlaGly-----GlyThrValLeuLysIleCysAsp 175
 DB 808 TTGCTACTTGAGAGATGAAGACATGATGATCTGCAATAACTTTGAAGATTACAGAT 867
 QY 176 PheGlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194

Pred. No.:	5,73e-27	Length:	2505
Score:	418.00	Matches:	120
Percent Similarity:	51.51%	Conservative:	68
Best Local Similarity:	32.88%	Mismatches:	138

Query Match:	13:87%	Indels:	40
DB:	10	Gaps:	13

Oy	31	11easpryrlvysguillegluvalglugluvalvalglvayrgglvalaPhgclvvalVal	50
Oy	11	leaspryrlvysguillegluvalglugluvalvalglvayrgglvalaPhgclvvalVal	50
Db	1372	CTTCAGGCTCTCAGAAATTTGATGTTCCATGTGAGATTATTTGGCTCAGGTTCTTTTGGGAAAGTA	1431
Oy	51	CyslvslatlvstTPaGlaValysAspValAlaIleLysGlnIleGlu-----	66
Db	1432	TATTAAGACGATTCACGAAATTAATAATAGTGGCTATTAACGTTATCGAGCCAAATCACTAC	1491
Oy	67	---SerglusergluarglvSalapheIleValGluLeuArgGlnLeuSerArgValAsn	85
Db	1492	TGCTCCAGGTGAGATGTGATGATTGTTTGGCCAGAGAGTGTCCATTCTCTCGACCTCAAT	1551
Oy	86	HisProAsnIleValLysLeuArglvArgValaCysLeuAsn-----ProValCysLeu	102
Db	1552	CATCCCTCGAGTAATTCAGATTGTTGGGGGCTGTTGAAATGATCCAGCCAGTTGGCAATT	1611
Oy	103	ValMetGlnUryAlaGluGluGlyGlySerLeuUryArgAsnValLeuHisGlyAlaGluProLeu	122
Db	1612	GTCACATCAATCAATATCAGAGGGGTTCTCTGTTCTCTCCCTTCATGCA-GCAGAAAGAGAT	1670
Oy	123	ProUryUryThrAlaAlaHisSalameSerUryCysLeuGlnCysSergln-GlyValAl	142
Db	1671	TCTTATTT-----GCAGTCTAAATTAATTAATTTGAGTAGATGTTGCCAAAGGATGGA	1724
Oy	142	ArgUryLeuHisSerMet---GlnProLysAlaLeuIleHisArgAspLeuLysProProAs	161
Db	1725	GTACTTTCACACACCTGCAGACAGCCA-----ATTATACATCTGACTTGAACAGTCAACA	1778
Oy	161	hLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAs	181
Db	1779	TATTTCTTCTCTATGAGGATGGGACGTGCTGTG---GTGGCAGATTTTGGAGATCAAGATT	1835
Oy	181	PIleGlnThr-----HisMetThrAsnAsnLysGlySerAlaAlaIleProMetAl	197
Db	1836	TCCTAACGTCCTCGATGACAGAACAAACATACAGAAAACAACCTGGGAACTCCGTTGATGTC	1895
Oy	197	ArgProGluValPhe---GlnGlySerAsnUrySerGlnUryCysAspValPheSerUryG	216
Db	1896	TCCTCAGGCTGTTCAOCCAGCTGCACTCGTATACACATCAAGAAGATGCTTCAGCTATGC	1955
Oy	216	YIleIleLeuUryGluValIleThrArgArgLysProPheAspGluIleGlyUryProAl	236
Db	1956	TCGTGTCTGTGGGAAATTTCTCATCTGCGCAAAATTCATTGCTCATCTCCAAOCCAGCGGC	2015
Oy	236	AlaPheArgIleMetCysPheAlaValHisAsnGlyUryArgProProLeuIleLysAsnLeuPr	256
Db	2016	TCGGCAGACAGACATGCTTACCAACCAC---ATCACAGCTTCCATTTGGGTATTCATTCC	2072
Oy	256	OlusProIleGluSerLeuMetThrArgCysUrySerUryAspProSerGlnArgProse	276
Db	2073	CAAGCCCAATCATCTCTGCTGATACAGAGGTGAAGCATGCTCCGAAGAAAGCCCGA	2132
Oy	276	rmMetGluGluIleValLysIleMetThrHisLeuMetArgUryPheProGluValAspG	296
Db	2133	ATTTTCTGAAGTTTCATGATGAAGTTAAGAGAGTGTCT-----TGCAACAT	2177
Oy	296	UryProLeuGlnUryProCysGlnUrySerAspGluGlyGlnSerAsnSerAlaThrSerTh	316
Db	2178	TGAGCTGATGTCCTCTGGA-----TCAAGTACACAGAG	2210
Oy	316	rgUrySerPheMetAspIleAlaSerThrAsn---ThrSerAsnLysSerAspThrAsnMe	335
Db	2211	TGGGTCTCTCACCCTTCTCTCTCTCTGATGTGCTCGTGAACCGGGGAGACCTGGCCG	2270
Oy	335	rGluGluValProAlaThrAsnAspThrIleLysArgLeuGlnSerUryLeuLeuLysAs	355
Db	2271	GAGTCATGTGGCAGCATTTAAGAAAGCTGTTTCGAAITGTGGAAITAGCTCTTAATGCAGAGCT	2330

Cy	355	ngmlalysgInglInserCIuserGIYargInserIuGlyalaserHISglYserse	375
		::: :::	
Db	2331	CTAGCTGCTTTTGTCCCAAGTGTCTGGACAAATTCC-----TTCAGAGGTCTTC	2381
Cy	375	rValGluserIeu	379
		::: :::	
Db	2382	TTTGGAGGAGATG	2394

```

10 RESULT
11 US-09-947-199-7
12 / Sequence 7, Application US/09947199
13 / Patent No. US2002012764A1
14 / GENERAL INFORMATION:
15 / APPLICANT: Raju, Jayaseelan
16 / TITLE OF INVENTION: NOVEL CARX PROTEIN AND NUCLEIC ACID MOLECULES AND USES
17 / TITLE OF INVENTION: THEREFOR
18 / FILE REFERENCE: MNI-068CP2
19 / CURRENT APPLICATION NUMBER: US/09/947,199
20 / CURRENT FILING DATE: 2001-09-05
21 / PRIOR APPLICATION NUMBER: 60/111,938
22 / PRIOR FILING DATE: 1998-12-11
23 / PRIOR APPLICATION NUMBER: 09/231,839
24 / PRIOR FILING DATE: 1999-04-14
25 / PRIOR APPLICATION NUMBER: 09/458,457
26 / PRIOR FILING DATE: 1999-12-10
27 / NUMBER OF SEQ ID NOS: 9
28 / SOFTWARE: PatentIn Ver. 2.0
29 / SEQ ID NO 7
30 / LENGTH: 3026
31 / TYPE: DNA
32 / ORGANISM: Rattus norvegicus
33 / FEATURE:
34 / NAME/KEY: CDS
35 / LOCATION: (61)..(2565)
36 /
37 US-09-947-199-7

```

Alignment Scores:	
Pred. No.:	7,41e-27
Score:	418.00
Percent Similarity:	44.33
Best Local Similarity:	28.238
Query Match:	13.878
DB:	10
Gaps:	
Length:	3026
Matches:	142
Conservative:	81
Mismatches:	161
Indels:	120
Gaps:	20

[illegible]

Qy	171	LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMet	186
Db	1807	GTG---GTGGCAGATTTTGGAGAAATCAAGATTTCGCAGTCCTCGATGAAGACAACATG	1863
Qy	187	ThrAsnAsnLysGlySerAlaAlaTTPMetAlaProGluValPhe--GluGlySerAsn	205
Db	1864	ACAAAGCAGCCAGGAACCTGCCCTGGATGGCCCTGAGGTGTTCCACAGTGCACAGAGA	1923
Qy	206	TyrSerGluLysCysAspValPheSerTrpGlyIlelleLeuTrpGluValIleThrArg	225
Db	1924	TACACATCAAGCGTATGTTCTTCAGTTACTCCCTGTGTCTGGAGCTCCTCAGTGA	1983
Qy	226	ArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsn	245
Db	1984	GAATTCCATTGCTCATCTCAAGCCAGCGCTGCAGCAGCAGATATGGCGTATCACCAC	2043
Qy	246	GlyThrArgProPheLeuLysAsnLeuProLysProIleGluSerLeuMetThrArg	265
Db	2044	---ATCAGACGGCCCATCGGCTATTCCATCCCAAGCCCATCTCATCCCTGCTGATACGG	2100
Qy	266	CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThr	285
Db	2101	GGCTGGAAATGCATGCTCTGAGGACACACAGAGTTCTCTGAAGTCTGTGAGAAATCGAG	2160
Qy	286	HisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTyrSer	305
Db	2161	GAGTGCCCTA-----TGCAATGTGGAGCTCATGCTCCACGCA-----	2196
Qy	306	AspGluGlnSerAsnSerAlaThrSerThrGlySerPheMetAspIleAlaSerThr	325
Db	2197	-----TCAGTAAACAGTGGCTCTCTG-----	2220
Qy	326	AsnThrSerAsnLysSerAspThrAsnMetGluGlnValProAlaThrAsnAspThrIle	345
Db	2221	TCACCTTCCTCTTCTCCGATTGCCTGTGAGCCGGGAGGCGCTGGCCGGAGCCACGTG	2280
Qy	346	LysArgGluLeuLysSerLysLeu--LeuLysAsnGlnAlaLysGlnGlnSerGluSerGly	364
Db	2281	GCAGCCTTACGAGCCGTTTTCAGTTGGAGTATGCCCTAAATGCAAGAGTCTCTATCTGGG	2340
Qy	365	ArgLeuSerLeuGlyAlaSerHis-----GlySerSerValGluSerLeu	379
Db	2341	TGTTCCCAAAGTGTGGAAACACACTCTAATCGGCGCCTGTCTTTGGAGGAGATG	2394

RESULT 12

US-09-938-842A-1014
; Sequence 1014, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; FILE REFERENCE: S1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1014
; LENGTH: 1662
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1014
Alignment Scores:

23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 105

DD 1301 AGOCHALCOGAMIGIOLALMIL

Db 1561 AGACATCCGAATGTGCTACTATTATGGGAG

```
QY 102 -----LeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
Db 1609 AAATCTCCATAATCATGGAATATATCCAGAGGGAGTCTCTCAAAATACCTTCAAT 1668
QY 119 Ala----GluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
Db 1669 ACGAATCAGCATG-----GACAAGAAACGCCGTTTAAGAATGGCCCTTGATGT 1719
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArgAspLeu 157
Db 1720 GCTAGGGGAATGAATTAATCTTACACCGCAGAAATCCG---CCAATTTGTACATAGACTTG 1776
QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
Db 1777 AAATCTTCCAACTACTCGTGGCAGCAAGACTGGAATGTC---AAGGTGGAGACTTTGGG 1833
QY 178 -----ThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySer 192
Db 1834 TTATCAAGTGGAGAACGCCAACCCTTCTTGAGTACT-----AAATCCGGGAAGGAACT 1887
QY 193 AlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspVal 212
Db 1888 CCGCAGTGGATGGCTCGTGAAGTTCTCAGAAGTGAACCTTCGAATGAGAAGTGTGATGTG 1947
QY 213 PheSerTrpGlyIleLeuTyrGluValIleThrArgArgLysProPheAspGluIle 232
Db 1948 TTCAGCTTTGGAGTCATCTTATGGAGCTAATGACACGTAGTACATGAGGACCGCTTTG 2007
QY 233 GlyGlyProAlaPheArgIleMetTrpAlaVal---HisAsnGlyThrArgProLeu 251
Db 2008 AAC-----TCTATTACAGTTGTGGAGTTGTGGTTTCATGATCGACGATTAGACTTA 2061
QY 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
Db 2062 CCTGAAGGATTAATCCCGGATCGCATCCATAATAACAGGATTGTTGGCAACATGATCCA 2121
QY 272 SerGlnArgProSerMetGluIleValLysIleMetThrHisLeuMetArgTyrPhe 291
Db 2122 GCAAAACGACCGTTCGAGGAATTAATTAGTCAGATGATGAGCCTGTTCCGCAAA--- 2178
QY 292 ProGlyAla 294
Db 2179 CCAGGGTCA 2187
```

RESULT 15

```
US-09-938-842A-1073
; Sequence 1073, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1073
; LENGTH: 2892
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1073
Alignment Scores:
```

```
Pred. No.: 1.81e-24 Length: 2892
Score: 390.00 Matches: 104
Percent Similarity: 51.08% Conservative: 62
Best Local Similarity: 32.00% Mismatches: 96
Query Match: 12.94% Indels: 64
DB: 9 Gaps: 14
US-09-830-144-2 (1-579) x US-09-938-842A-1073 (1-2892)
QY 22 ProSerGlnValLeuAsnPheGluGluLeuAspTyrLysGluIleGluValGluVal 41
Db 1963 CTTCTTTGGCATCAGATTGGCTGAGGTCTCATGGAATGAATTCATATCAATATCAAAGCGT 2022
QY 42 ValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspValAla 61
Db 2033 GTGGGTGCTGGATCATTTGGAACTGTTTCATCGTCTGAGTGGCATGATGATGTTGCT 2082
QY 62 IleLys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGlu 77
Db 2083 GTCAAGATTTTGTCTATTTCAGATTTCATGATGACCAATTCAGAGAAATTTCTCAGAGAG 2142
QY 78 LeuArgGln-----LeuSerArgValAsnHisProAsnIleValLysLeu 92
Db 2143 GTATGTAAGCAAGCGGTGCTATAATGAACGCTGCTCGTCAACCAATGTTGTCT-CTT 2201
QY 93 TyrGlyAlaCys-----LeuAsnProVal-CysLeuValMetGluTyrAlaGluGly 110
Db 2202 CATGGGTGCTGTGACAGAGCGACCCCGGTATCAATAATAACAGAAATATTGTCACAGAGG 2261
QY 110 ySerLeuTyrAsnValLeuHisGlyAlaGlu----- 120
Db 2262 CAGCTTTTTCGCTTATCCATAGCCAGCTTCTGGGAGTTGCTAGATCAGAGGAGGAG 2321
QY 121 -----ProLeuProTyrTyrThrAlaAlaHisAlaMe 131
Db 2322 GCTACGTATGTCATTGGATGTTGTTGCTATTCCCATTTAT----- 2364
QY 131 tSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAl 151
Db 2365 -----GCCAAGGGGCTCACTACTACTACTGCTCTTAACTCT-CC 2402
QY 151 aLeuIleHisArgAspLeuLysProCAsnLeuLeuValAlaGlyGlyThrValle 171
Db 2403 TGTAGTGCATTGGGACCTGAAATCTCCAAATCTACTGTTGATGAAGAACTGGACAGTG-- 2460
QY 171 uLysIleCysAspPheGly---ThrAlaCysAspIleGlnThrHisMetThrAsnLys 190
Db 2461 -AAGGTTTTCGATTTTGGACTTTTCAAGATTCAAGGCAACACACTTTCATACCATCAAAATC 2519
QY 190 s-----GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSer 208
Db 2520 TGTTCAGGAACACCTGAGTGGATGGCTCCAGAGTTTCTTAGAGGGGAAACCGACAAACGA 2579
QY 208 uLysCysAspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysPr 228
Db 2580 GAATCAGATGTTTACAGTTTCGGAGTAGTCTTATGGAGTTGATTAATTTGCAACAGCC 2639
QY 228 oPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrAr 248
Db 2640 TTGGAATGGACTC---AGTCTCTGCT---CAGGTGGTTGGAGCAGTTGTCATTCCAGAACAG 2693
QY 248 g-----ProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetTh 264
Db 2694 CGGCTTATAATTTCTTCCC-----AACACCTCTCCGCTTTTGTGATCTCTTAATGGA 2744
QY 264 rArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMe 284
Db 2745 AGCTTGTGGGCAGATGAGCGCTCTCAGCGCCAGCACTTTGGTAGTAGTAGTGACACATT 2804
QY 284 tThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCysGlnTy 304
Db 2805 GAAGAAGCTACTAAAG-----TCTCCGGTGCAGCTGATCCAAATGGG 2846
```

Qy 304 rSerApGInGly 308
|||::|||
Db 2847 TGGAGCAAGGG 2859

Search completed: December 10, 2002, 02:09:05
Job time : 111.402 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 02:14:40 ; Search time 87 Seconds
(without alignments)
464.080 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303
Perfect score: 1615
Sequence: 1 MSTASASSSSSSSAGEWIE.....MTHLMRFPADPELQPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	567	20	AA28998 Human TGF-beta act
2	1615	100.0	579	18	AA27093 Human transforming
3	1615	100.0	579	20	AA28996 Human TGF-beta act
4	1615	100.0	579	20	AA209542 Human TAK1 protein
5	1615	100.0	579	21	AA291000 Human TAK-1 protei
6	1615	100.0	579	23	AB85033 Pain regulated pro
7	1615	100.0	590	20	AA209547 Human TAK1-6xHis p
8	1615	100.0	606	20	AA28997 Human TGF-beta act
9	1609	99.6	579	18	AA27092 Mouse transforming
10	813	50.3	678	22	AB58061 Drosophila melanog

11	505	31.3	252	22	AB80985 Drosophila melanog
12	434.5	26.9	367	21	AA32053 Arabidopsis thalia
13	434.5	26.9	369	21	AA32052 Arabidopsis thalia
14	434.5	26.9	407	21	AA32051 Arabidopsis thalia
15	434	26.9	349	22	AA75571 Human colon cancer
16	434	26.9	369	21	AA32172 Arabidopsis thalia
17	434	26.9	374	21	AA32171 Arabidopsis thalia
18	434	26.9	412	21	AA32170 Arabidopsis thalia
19	434	26.9	473	22	AA25322 Human protein sequ
20	433	26.8	455	21	AA18657 A human regulator
21	433	26.8	455	21	AA783278 Human survival reg
22	433	26.8	455	21	AA784321 A human cardiovascular
23	433	26.8	800	22	AA31957 Human TGF-beta rec
24	433	26.8	800	22	AA35673 Human protein kina
25	433	26.8	1046	22	AA311775 Human kinase (PKIN
26	432.5	26.8	1021	23	ABP61000 Novel human protei
27	432	26.7	719	22	AA35553 Human protein kina
28	432	26.7	1036	23	AB80993 Novel human protei
29	431	26.7	1097	23	AA321717 Human PKIN-12 prot
30	427.5	26.5	341	21	AA325600 Arabidopsis thalia
31	427.5	26.5	391	21	AA325599 Arabidopsis thalia
32	427.5	26.5	1020	22	AB358999 Drosophila melanog
33	419	25.9	847	23	AA322763 Human mitogen acti
34	411.5	25.5	589	21	AA345984 Arabidopsis thalia
35	411.5	25.5	732	21	AA345983 Arabidopsis thalia
36	411.5	25.5	760	21	AA345982 Arabidopsis thalia
37	410	25.4	835	21	AA345981 Human CARD (Cardia
38	410	25.4	835	21	AA345980 Rat CARD (Cardiac
39	410	25.4	835	22	AA345979 Novel protein kina
40	410	25.4	928	22	AB316533 Novel human diagno
41	404.5	25.0	319	21	AA325601 Arabidopsis thalia
42	401	24.8	859	16	AA32886 Human leucine-zipp
43	401	24.8	888	23	AA31227 Human leucine-zipp
44	400	24.8	888	23	AB37049 Mouse ischaemic co
45	399	24.7	977	22	AB371694 Drosophila melanog

ALIGNMENTS

RESULT 1
AA28998 ID AA28998 standard; Protein; 567 AA.
XX
AC AA28998;
XX
DT 29-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1c amino acid sequence.
XX
KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c.
XX
OS Homo sapiens.
XX
PN WO9940202-A1.
XX
PD 12-AUG-1999.
XX
PF 02-FEB-1999; 99WO-0P00422.
XX
PR 30-OCT-1998; 98UP-0309316.
XX
PR 06-FEB-1998; 98UP-0026003.
XX
PA (TANA) TANABE SEIYAKU CO.
XX
PI Haegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI, 1999-494298/41.
XX DR N-PSDB; AAX99698.
XX

PT Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases
 XX
 PS Examples; Page 43-46; 49pp; Japanese.
 XX
 CC The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the amino acid sequence of human TAK1c
 CC (hTAK1c) protein.
 XX
 XX Sequence 567 AA;
 SQ

Query Match 100.0%; Score 1615; DB 20; Length 567;
 Best Local Similarity 100.0%; Pred. No. 2e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||
 Db 1 MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||

QY 61 AIKQIESESEKAFIVELQSLRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
 |||||
 Db 61 AIKQIESESEKAFIVELQSLRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
 |||||

QY 121 PLPYTTAAHAMSCLQCSOGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||
 Db 121 PLPYTTAAHAMSCLQCSOGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||
 Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||

QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSDPSQSPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||
 Db 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSDPSQSPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||

QY 301 PCQ 303
 |||||
 Db 301 PCQ 303

RESULT 2
 AAW27093
 ID AAW27093 standard; Protein; 579 AA.
 XX
 AC AAW27093;
 XX
 DT 19-NOV-1997 (first entry)
 XX
 DE Human transforming growth factor-beta activated kinase TAK-1.
 XX
 DE TGF-beta; signal transduction; TGF-beta activated kinase;
 KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
 KW protein kinase.
 XX
 OS Homo sapiens.
 XX
 XX JP09163990-A.
 PN
 XX 24-JUN-1997.
 XX
 XX 27-SEP-1996; 96JP-0256747.
 XX
 XX 24-JUL-1996; 96US-0685625.
 PR 29-SEP-1995; 95JP-0253549.
 XX
 XX (CHUS) CHUGAI PHARM CO LTD.
 PA (UENO/) UENO N.

XX WPI: 1997-380171/35.
 DR N-PSDB; AAT85095.
 XX
 PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
 PT - useful for studying the TGF-beta signal transduction system
 XX
 PS Claim 15; Page 13-15; 20pp; Japanese.
 XX
 CC The present sequence represents human transforming growth factor-beta
 CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
 CC TAK-1 protein which is involved in the TGF-beta family signal
 CC transduction system. TAK-1, also known as activator of MAPK kinase
 CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
 CC morphogenetic protein (BMP) and activates MAPK kinase by
 CC phosphorylation.
 XX
 XX Sequence 579 AA;
 SQ

Query Match 100.0%; Score 1615; DB 18; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.1e-170;
 Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||
 Db 1 MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
 |||||

QY 61 AIKQIESESEKAFIVELQSLRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
 |||||
 Db 61 AIKQIESESEKAFIVELQSLRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
 |||||

QY 121 PLPYTTAAHAMSCLQCSOGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||
 Db 121 PLPYTTAAHAMSCLQCSOGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
 |||||

QY 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||
 Db 181 DIQHTMTNNKGSAAWMAPEVFEKSNYSEKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIM 240
 |||||

QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSDPSQSPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||
 Db 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSDPSQSPSMEEIVKIMTHLMRYFFGADEPLQY 300
 |||||

QY 301 PCQ 303
 |||||
 Db 301 PCQ 303

RESULT 3
 AAY28996
 ID AAY28996 standard; Protein; 579 AA.
 XX
 AC AAY28996;
 XX
 DT 29-OCT-1999 (first entry)
 XX
 DE Human TGF-beta activated kinase (TAK) la amino acid sequence.
 XX
 DE Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1a.
 XX
 OS Homo sapiens.
 XX
 XX WO9940202-A1.
 PN
 XX 12-AUG-1999.
 XX
 XX 02-FEB-1999; 99WO-JP00422.
 PR 30-OCT-1998; 98JP-0309316.
 XX
 XX 06-FEB-1998; 98JP-0026003.


```

XX (TANA ) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
XX DR N-PSDB; AAX99696.
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX
XX Examples; Page 35-39; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NPKB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intracellular diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the amino acid sequence of human TAK1a
XX (hTAK1a) protein.
XX
XX Sequence 579 AA;
SQ
Query Match 100.0%; Score 1615; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.1e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTASASSSSSSAGEMTAPSOVLNFEEDYKEIEVEEVVGAGFGVCXAKMRKADV 60
DB 1 MSTASASSSSSSAGEMTAPSOVLNFEEDYKEIEVEEVVGAGFGVCXAKMRKADV 60
QY 61 AIKQIESSEKRAFIIVELRQLSRVNHPNIVKLYGACINPVCVLMVEYAGGSLYNVLHGAE 120
DB 61 AIKQIESSEKRAFIIVELRQLSRVNHPNIVKLYGACINPVCVLMVEYAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
QY 181 DIQHTMTNNKGSAAAMWAPVEFEGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARFRIM 240
DB 181 DIQHTMTNNKGSAAAMWAPVEFEGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARFRIM 240
QY 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQ 303
DB 301 PCQ 303

```

RESULT 4
AA909542
ID AA909542 standard; Protein; 579 AA.

AC AA909542;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAK1 protein.
XX
KM Human; TAB1; TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 183..1922
XX /*tag= a

```

PN W09921010-A1.
XX
XX 29-APR-1999.
XX
XX 22-OCT-1998; 98WO-JP04796.
XX
XX 22-OCT-1997; 97JP-0290188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Ohtomo T, Ono K, Tsuchiya M;
XX
XX WPI; 1999-312645/26.
XX DR N-PSDB; AAX56279.
XX
XX Screening for TGF-beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
XX Claim 4; Page 155-157; 195pp; Japanese.
XX
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transmission inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors or activators, or cell proliferation prevention inhibitors or
XX activators, or monocyte migration inhibitors or activators, or
XX physiological activity induction inhibitors or activators, or
XX immunosuppression inhibitors or activators, or amyloid beta protein
XX precipitation inhibitors or activators, and such substances can also be
XX inhibitors of the TAK1 polypeptide function, particularly kinase
XX activity. The present sequence represents human TAK1.
XX
XX Sequence 579 AA;
SQ
Query Match 100.0%; Score 1615; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.1e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTASASSSSSSAGEMTAPSOVLNFEEDYKEIEVEEVVGAGFGVCXAKMRKADV 60
DB 1 MSTASASSSSSSAGEMTAPSOVLNFEEDYKEIEVEEVVGAGFGVCXAKMRKADV 60
QY 61 AIKQIESSEKRAFIIVELRQLSRVNHPNIVKLYGACINPVCVLMVEYAGGSLYNVLHGAE 120
DB 61 AIKQIESSEKRAFIIVELRQLSRVNHPNIVKLYGACINPVCVLMVEYAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
QY 181 DIQHTMTNNKGSAAAMWAPVEFEGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARFRIM 240
DB 181 DIQHTMTNNKGSAAAMWAPVEFEGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPARFRIM 240
QY 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIIVKIMTHLMRYFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLKPRIEISLMTKCSKDPSPRSMEIIVKIMTHLMRYFPGADEPLQY 300
QY 301 PCQ 303
DB 301 PCQ 303

```

RESULT 5
AA91000
ID AA91000 standard; Protein; 579 AA.

AC AA91000;
XX

```
DT 04-SEP-2000 (first entry)
XX Human TAK-1 protein sequence SEQ ID NO:2.
XX
XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression.
XX
OS Homo sapiens.
XX WO200023610-A1.
XX
XX 27-APR-2000.
XX
XX 21-OCT-1999; 99WO-JP05817.
XX
XX 21-OCT-1998; 98JP-0299962.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
PI WPI; 2000-339707/29.
XX
XX N-PSDB; AAA39105.
XX
XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents
XX
XX Example 1; Page 80-84; 100pp; Japanese.
XX
XX The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug
CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence represents human TAK-1, which is used in the exemplification of
CC the present invention.
XX
XX SQ Sequence 579 AA;
XX
XX Query Match 100.0%; Score 1615; DB 21; Length 579;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-170;
XX Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
XX
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
XX
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLKPNLLLVAGGTVLKICDFTGAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLKPNLLLVAGGTVLKICDFTGAC 180
XX
QY 181 DIOTHTNNKGSAAWMAPEVFGNSYSEKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
DB 181 DIOTHTNNKGSAAWMAPEVFGNSYSEKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
XX
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLOY 300
XX
QY 301 PCQ 303
DB 301 PCQ 303
XX
XX RESULT 6
XX ID ABB85033 standard; Protein; 579 AA.
XX
XX AC ABB85033;
XX
XX 16-MAY-2002 (first entry)
XX
XX Pain regulated protein sequence 28.
XX
XX Pain; analgesic; gene therapy; neurological disorder;
KW neurodegenerative disease.
XX
XX Homo sapiens.
XX
XX WO200212338-A2.
XX
XX 14-FEB-2002.
XX
XX 03-AUG-2001; 2001WO-EP09011.
XX
XX 03-AUG-2000; 2000DE-1037759.
XX
XX (CHEF ) GRUENENTHAL GMBH.
XX
XX Gillen C, Wetzels I, Wnendt S, Weihe E, Schaefer MK;
PI WPI; 2002-257469/30.
XX
XX N-PSDB; ABL88437.
XX
XX Identifying pain-regulating compounds, useful for treating chronic pain
PT and for diagnosis, by measuring binding of compounds to specific
PT peptides and proteins
XX
XX Claim 1; Fig 44; 213pp; German.
XX
XX The invention relates to identifying pain-regulating substances (A)
CC comprises (i) incubating a test substance with a cell (or preparation
CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying
CC pain-regulating substances (A) with analgesic activity. (A) along with
CC nucleic acid (ABL88411-ABL88441) that interact with (A); (B); vectors containing the
CC nucleic acid; antibodies against (B); cells that express (B) and agents
CC that bind to (B), are all useful for treating pain, particularly chronic
CC pain, including use in gene therapy. The same materials can also be used
CC for diagnosis, e.g. of neurological and neurodegenerative diseases. The
CC present sequence is that of a polypeptide of the invention.
XX
XX SQ Sequence 579 AA;
XX
XX Query Match 100.0%; Score 1615; DB 23; Length 579;
XX Best Local Similarity 100.0%; Pred. No. 2.1e-170;
XX Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVGGRGAFGVVCKAKWRAKDV 60
XX
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
XX
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLKPNLLLVAGGTVLKICDFTGAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLKPNLLLVAGGTVLKICDFTGAC 180
XX
QY 181 DIOTHTNNKGSAAWMAPEVFGNSYSEKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
DB 181 DIOTHTNNKGSAAWMAPEVFGNSYSEKCDVPSWGIILWEVITTRKPFDEICGGPAFRIM 240
XX
QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPRSMEEIVKIMTHLMRYFPGADEPLOY 300
XX
```

```

QY 181 DIQHTMTNNKGSAAWMAPEVFEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAWMAPEVFEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
QY 241 WAWHNGTRPPLIKNLKPPIESLMTRCWSDPSQRPSPMEIVKIMTHLMRYPPGADEPIQY 300
DB 241 WAWHNGTRPPLIKNLKPPIESLMTRCWSDPSQRPSPMEIVKIMTHLMRYPPGADEPIQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 7
AA09547
ID AA09547 standard; Protein; 590 AA.
AC AA09547;
XX 21-JUL-1999 (first entry)
DE Human TAK1-6xHis protein.
XX Human TAK1; screening; inhibition; TGF-beta;
XX transforming growth factor beta.
XX Homo sapiens.
XX Synthetic.
XX WO9921010-A1.
XX 29-APR-1999.
XX 22-OCT-1998; 98WO-JP04796.
XX 22-OCT-1997; 97JP-0290188.
XX (CHUS ) CHUGAI SEIYAKU KK.
XX Ohtomo T, Ono K, Tsuchiya M;
XX WPI; 1999-312645/26.
XX N-PSDB; AAX56285.
XX Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX Example 1; Page 171-174; 195pp; Japanese.
XX A method has been developed for screening for substances which inhibit
XX the binding of TAK1 polypeptide to TAB1 polypeptide. The method
XX comprises: (a) contacting the polypeptide in the presence of a sample;
XX and (b) detecting the amount of bound polypeptide, in which the sample
XX can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
XX growth factor (TGF)-beta inhibitory substances can be used in drugs for
XX indications e.g. as TGF-beta signal transduction inhibitors or
XX activators, or extracellular matrix protein production enhancement
XX inhibitors, or monocycle migration inhibitors or activators, or
XX physiological activity induction inhibitors or activators, or
XX immunosuppression inhibitors or activators, or amyloid beta protein
XX precipitation inhibitors or activators, and such substances can also be
XX inhibitors of the TAK1 polypeptide function, particularly kinase
XX activity. The present sequence represents TAK1-6xHis from an example of
XX the present invention.
XX Sequence 590 AA;
XX
Query Match 100.0%; Score 1615; DB 20; Length 590;
Best Local Similarity 100.0%; Pred. No. 2,2e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSTAASASSSSSGAGMIAPSVQVNFEBIDYKEIEVEEVGARGAVGCKAKMRAKV 60

```

```

DB 1 MSTAASASSSSSGAGMIAPSVQVNFEBIDYKEIEVEEVGARGAVGCKAKMRAKV 60
QY 61 AIKOISESEFRKAFIYELIQLSRVNHPNIVKYGACLVNCLVMEYABGSLYNNVHGA 120
DB 61 AIKOISESEFRKAFIYELIQLSRVNHPNIVKYGACLVNCLVMEYABGSLYNNVHGA 120
QY 121 PLPYTAAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKRPNNLLVAGGVLYKICPGTAC 180
DB 121 PLPYTAAAHAMSWCLOCSQGVAYLHSMOPKALIHRLKRPNNLLVAGGVLYKICPGTAC 180
QY 181 DIQHTMTNNKGSAAWMAPEVFEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
DB 181 DIQHTMTNNKGSAAWMAPEVFEKDVFSWGIILMEVITRRKPEDEIGGPAFRIM 240
QY 241 WAWHNGTRPPLIKNLKPPIESLMTRCWSDPSQRPSPMEIVKIMTHLMRYPPGADEPIQY 300
DB 241 WAWHNGTRPPLIKNLKPPIESLMTRCWSDPSQRPSPMEIVKIMTHLMRYPPGADEPIQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 8
AA28997
ID AA28997 standard; Protein; 606 AA.
XX AA28997;
XX 29-OCT-1999 (first entry)
XX Human TGF-beta activated kinase (TAK) 1b amino acid sequence.
XX Human TGF-beta activated kinase (TAK) 1b amino acid sequence.
XX Nuclear factor kappa B, NF-kB, inhibitor; TGF-beta activated kinase 1;
XX TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
XX intractable disease; atrophic dermatitis; psoriasis; viral infection;
XX endotoxin shock; septicemia; human; hTAK1b.
XX Homo sapiens.
XX WO9940202-A1.
XX 12-AUG-1999.
XX 02-FEB-1999; 99WO-JP00422.
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.
XX (TANA ) TANABE SEIYAKU CO.
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX WPI; 1999-494298/41.
XX N-PSDB; AAX93697.
XX Nuclear factor kappa B activation inhibitors, useful as preventives
XX for, e.g. autoimmune diseases
XX Examples; Page 39-43; 49pp; Japanese.
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intractable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the amino acid sequence of human TAK1b
XX (hTAK1b) protein.
XX Sequence 606 AA;
XX

```

Query Match 100.0%; Score 1615; DB 20; Length 606;
Best Local Similarity 100.0%; Pred. No. 2.3e-170;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIQTHMTNNKGSAAWMAPEVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADPELQY 300
DB 241 WAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADPELQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 9
AAW27092
ID AAW27092 standard; Protein; 579 AA.
XX
AC AAW27092;
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1.
XX
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Mus musculus.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO/) UENO N.
XX
WPI; 1997-380171/35.
DR N-PSDB; AAT85094.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
PS Claim 14; Page 10-12; 20pp; Japanese.
XX
CC The present sequence represents mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.

XX
SQ Sequence 579 AA;
Query Match 99.6%; Score 1609; DB 18; Length 579;
Best Local Similarity 99.7%; Pred. No. 9.8e-170;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTASASSSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIQTHMTNNKGSAAWMAPEVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADPELQY 300
DB 241 WAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQSPSMEEIVKIMTHLMRYFFGADPELQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 10
ABB58061
ID ABB58061 standard; Protein; 678 AA.
XX
AC ABB58061;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 975.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
Venter JC, Adams M, Li PWD, Myers EW;
WPI; 2001-656860/75.
DR N-PSDB; ABL02164.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 678 AA;
Query Match 50.3%; Score 813; DB 22; Length 678;
Best Local Similarity 54.0%; Pred. No. 5.4e-81;
Matches 154; Conservative 45; Mismatches 82; Indels 4; Gaps 3;
QY 18 MIEAPSOVLNEEIDYKEIEVEEVGAGFVCKAKRKADVAIKQIESSEKAFIVE 77
DB 1 MATASLDLQAAYVDFSBITLREKVGHSYGVCKAVMRDLVAVKFPFASAGKDIKE 60
QY 78 LROLSRVNHPVIVLYG--ACLNVCLVMEYAEAGSLYVNLHGAEPLPYTTAAHMSWCL 135
DB 61 VKQLSRVGHPPNIIALHGISSYQOATYIMEFAEGSLHNFJLHG-KVNPAYSLAHAMSMAR 119
QY 136 QCSGAVLIHSMOPKALIHRLDKPNNLLVAGGTVLKICDGTACDIOTHTNNKSAAM 195
DB 120 QCAAGLAIYLAHMTKPLIHRDYKPLNLLTNKGNLKI CDFTVADKSTMTNNGSAAM 179
QY 196 MAPEVFGSNYSKCDVFSWGIIIMEVITRRKPPDEIGPAFRIMAWHNGTRPPLIKNL 255
DB 180 MAPEVFGSKYTEKCDIFSMAIVLMEVLSRKQPKGIDN-AYTIQWKLTYGGERPPLITTC 238
QY 256 PKPIESLMTRCWSDPGORPSMEIIVKIMTHLMRYFPQADEPLPY 300
DB 239 PKRIEDLMTACWKTVPEDRPMQYIVGMHEIVKDYTGADKALEY 283
RESULT 11
ABB60985
ID ABB60985 standard; Protein; 252 AA.
XX
AC ABB60985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 9747.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KM pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX N-PDB; ABL05088.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Disclosure; SEQ ID NO 9747; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL10511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (AB857737-AB872072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 252 AA;
Query Match 31.3%; Score 505; DB 22; Length 252;
Best Local Similarity 40.6%; Pred. No. 2.2e-47;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;
QY 29 EEIDYKEIEVEEVGAGFVCKAKRKADVAIKQIESSEKAFIVELRQLSRVNHPN 88
DB 6 EGVPEEIEIQTKELIGTFGYSYRAVMNRRIALKRIRGCGDKIEREIVQLTRASHVN 65
QY 89 IVKLYGACLNPCV--LVMEYAGSLYVNLHGAEPLPYTTAAHMSWCLQSGVAYIHS 146
DB 66 IYELIGTSRHEGCAULLMEFYDGSLSFLLH-AKSPSYSHAHAFNWAHOIAQGIAYLHG 124
QY 147 MQPKALIHRLDKPNNLLVAGGTVLKICDGTACDIOTHTNNKSAAMMAPE----- 199
DB 125 MQPKAVIHHDIPNLTLTLCEKGLKLIKIDFTGVVDLSOSICNACTCYKXAEVRELPDF 184
QY 200 -----VFEGSNYSKCDVFSWGIIIMEVITRRKPPDEIGPAFRIMAWH 244
DB 185 KSNRIIINOPTFOKVLQGNKPKDCKDVSMATFWELISRKEPEQ-NTLFELYMALN 243
QY 245 NGTR 248
DB 244 EGR 247
RESULT 12
AAG32053
ID AAG32053 standard; Protein; 367 AA.
XX
AC AAG32053;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 38597.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
XX
PR 05-MAR-1999; 99US-0123180.
XX
PR 09-MAR-1999; 99US-0123548.
XX
PR 23-MAR-1999; 99US-0125788.
XX
PR 25-MAR-1999; 99US-0126284.
XX
PR 29-MAR-1999; 99US-0126785.
XX
PR 01-APR-1999; 99US-0127462.
XX
PR 06-APR-1999; 99US-0128234.
XX
PR 08-APR-1999; 99US-0128714.
XX
PR 16-APR-1999; 99US-0129845.
XX
PR 19-APR-1999; 99US-0130077.
XX
PR 21-APR-1999; 99US-0130449.
XX
PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132488.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 08-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 26-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.

PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 434.5; DB 21; Length 367;
Best Local Similarity 34.8%; Pred. No. 2,8e-39;
Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;

QY 12 SSSAGEMIEAPSOY-----LNFEELIDYKEIEVEEVRGGA 46
DB 37 SLASYGVFRGRTHALNDALQALMDTRYPTREGLTNDWTIDRKINMGPAFAQGA 96
QY 47 FGVCCKAKRAKDVAIKQIE--SESEKRA-----FVLELRQLSRVNHPNIVKLYGACLN 98
DB 97 FGKLYKGTYNEDVAIKILERPENSPEKAQFMEOQFOEVSMLANLHPNIVPRIGACRK 156
QY 99 PV--CLMVEYAEGLNVL---HGAEPLPYTAAHAMSCLOCSGVAYLHSMQKAL 152
DB 157 PMWVCILTEYVAKGSGVQFLTRQNRVPLKL-----AVKQALDVAGMAVYHG--RNF 208
QY 153 IHRDLKPNLLVAGGTVLKICDFGTA-CDIQTH-MTNNGSAAWMAPEVEGSGNYSEKC 210
DB 209 IHRDLKSNLISADKSI-KIADFGVARIEVQTEGMPETICTYMMAPBMTQHRAYNQKV 267
QY 211 DVFSWGIILWEVITRRKPFDEIGG--PAFRIMAVHNGTRPPLIKNLPKPIESIMTRCWS 268
DB 268 DVYSGFYLWELITGLLPFQNMVAQAFAV---VNRGVRTVENDCLPYLSDIMTRCWD 324
QY 269 KDPQRPSPMEIIVKIM 284
DB 325 ANPEVRPCFEVVKIL 340

RESULT 13
AAG32052
ID AAG32052 standard; Protein; 369 AA.

XX AC AAG32052;

DT 17-OCT-2000 (first entry)

XX DX Arabidopsis thaliana protein fragment SEQ ID NO: 38596.

XX KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX OS Arabidopsis thaliana.

XX PN EPI033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0136782.
PR 03-JUN-1999; 99US-0137222.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137502.
PR 08-JUN-1999; 99US-0137724.
PR 10-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140355.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0144806.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 26.9%; Score 434.5; DB 21; Length 369;
Best Local Similarity 34.8%; Pred. No. 2.9e-39;

Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;

QY 12 SSSAGEMIEAPSQV-----LNFE--IDYKEIEVEEVVGRGA 46
DB 39 SLVGSQSVFRPRVTHALDDALQAALMDPRYPTEGLTNYDEWTIDLRKLNMGPAFAQGA 98
QY 47 FGVCYKAKWRAKVAIQIE--SESEKA-----FIVELRQLSRVNHVNIKLYGACLN 98
DB 99 FGKLYKGTNGEDVAIKILERPENSPEKAQFMEQQEQEVSMLANLKHFNIVRFIGACRK 158
QY 99 PV--CLVMEYAEAGGSLYNVL---HGAELPYVYTAHAHMSWCLQCSQGVAYLHSMQPKAL 152
DB 159 PMWCIVTEYAKGSGVRQFLTRRQNRAPVPLK-----AVKQALDVARGMAYVHG---RNF 210
QY 153 IHRDLKPPNLLVAGGTVLKICDFGTA-CDIQTH-MTNKNGSAAWMAPEVFECSNYSEK 210
DB 211 IHRDLKSDNLLISADKSI-KIADFGVARIIVQTEGTPETGTYRWAPENIQHAYNQKV 269
QY 211 DVPSWGIILWEVITRRKPPDEIG--PAPRIMWAVHNGTRPPLIKNLKPDIKPLBLSMTRCWS 268
DB 270 DVYSFGIVLWELITGLLPFQNMATAVQAFAV---VNRGVRPTVPNDCLPVLSDIMTRCWD 326
QY 269 KDRSQRPSEMEIVKIM 284
DB 327 ANPEVRFCFVEVVKLL 342

RESULT 14

AAG32051
ID AAG32051 standard; Protein; 407 AA.

XX AAG32051;

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 38595.
DE

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132488.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0143977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151068.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.

	: : : : : :	
Db	365 ANPEVRPCFVGVVKLL 380	
	RESULT 15	
	AAG75571	
ID	AAG75571 standard; Protein; 349 AA.	
XX	AAAG75571;	
AC	AAG75571;	
XX	03-SEP-2001 (first entry)	
DT	Human colon cancer antigen protein SEQ ID NO:6335.	
DE	Human; colon cancer; colon cancer antigen; diagnosis; detection;	
KW	colorectal carcinoma.	
XX	Homo sapiens.	
OS	WO200122920-A2.	
XX	05-APR-2001.	
PN	28-SEP-2000; 2000WO-US26524.	
XX	29-SEP-1999; 99US-0157137.	
PF	03-NOV-1999; 99US-0163280.	
PR	(HUMA-) HUMAN GENOME SCI INC.	
XX	Ruben SM, Barash SC, Birse CE, Rosen CA;	
PI	WPI; 2001-235357/24.	
XX	N-PSDB; AAH34976.	
DR	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,	
PT	useful for preventing, diagnosing and/or treating colorectal cancers -	
XX	Claim 11; Page 7789-7790; 9803pp; English.	
PS	AAH32943 to AAH37195 and AAG73514 to AAG77798 represent human colon	
CC	cancer-associated nucleic acid molecules (N) and proteins (P), where	
CC	the proteins are collectively known as colon cancer antigens. The colon	
CC	cancer antigens have cytostatic activity and can be used in gene	
CC	therapy and vaccine production. N and P may be used in the prevention,	
CC	diagnosis and treatment of diseases associated with inappropriate P	
CC	expression. For example, N and P may be used to treat disorders	
CC	associated with decreased expression by rectifying mutations or deletions	
CC	in a patient's genome that affect the activity of P by expressing	
CC	inactive proteins or to supplement the patients own production of P.	
CC	Additionally, N may be used to produce the colon cancer-associated Ps,	
CC	by inserting the nucleic acids into a host cell and culturing the cell	
CC	to express the proteins. N and P can be used in the prevention, diagnosis	
CC	and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204	
CC	and AAB77789 represent sequences used in the exemplification of the	
CC	present invention.	
CC	N.B. Pages 666 to 682 and page 7053 of the sequence listing were	
CC	missing at time of publication, meaning no sequences are present for	
CC	SEQ ID NO:1027 to 1052, 7921 and 7922.	
XX	Sequence 349 AA;	
SQ	Query Match 26.9%; Score 434.5; DB 21; Length 407;	
	Best Local Similarity 34.8%; Pred. No. 3.4e-39; Indels 55; Gaps 13;	
	Matches 110; Conservative 54; Mismatches 97; Indels 55; Gaps 13;	
Qy	12 SSSAGEIEAPSOV-----LNFEI--IDYKEIEVEVVGRGA 46	
Db	77 SLVGQSVPFRGRTVALNDALQAALMDRYPTTEGLTYDWTIDLRLKLMGPAPAQA 136	
Qy	47 FGVCCKAKRAKVAIKQIB--SESERKA-----FIVELQLSRVNHNVKLKGACLN 98	
Db	137 FGKLYKGTNGEDVIAIKILERPENSPEKAFMEQQFOQEVSVMIANLUKHPIVRFIGACKR 196	
Qy	99 PV--CLNVEYAEGSLXNLV----HGAEPLPYTAAHAHSWCILQCSCGVAYLHSMOPKAL 152	
Db	197 PMWCICITEYAKGSVRFQFLTRQNRAVLPLK-----AVKQALDVARGMAYVHG---RNF 248	
Qy	153 IHRDLKPNNLLL VAGGTVLKICDPGTA-CDIQTH-MTNKKSAAWNAPEVFEGSNYSSEKC 210	
Db	249 IHRDLKSDNLLISADKSI-KIADFVARIETVQEGMTPTGTGYRWMAPEMIQHRAINQKV 307	
Qy	211 DVFSWGIIILEVITRRKFPDFEIGG--PAFRIMVAHVHTGTPPLIKLPKPIESLMTRCS 268	
Db	308 DVSFGVILVELIITGLIPFQNMWTAQAAFAV---VNRGVPTVPNDCLPVLSIDIMTRCD 364	
Qy	269 KDPSPRSMEEIVKIM 284	

```
QY 131 MSWCLQCSQGVAYLHSMQPKALIHRLKFPNLLVAGSTVLKICDEGTACDIOTHTNKK 190
Db 164 MTWATDVAKGMHYLHMEAPVKYIHRDLKSRNVIAADG-VLKI CDFG-ASRFHHTTHMS 221
QY 191 --GSAAMMAPEVEGSGNYSSEKCDVFSWGIIIMWEVITRRKFPDEIGGPAFRIMW-AVHNGT 247
Db 222 LVGTFFPMAPEVIQSLPVSETCDTYSYGVVIMEMLTREVFPKGLEG--LQYAMLVEKNE 279
QY 248 RPPLINKLPKPIESLMTRCWSKDPGQPSMEIIVKIM 284
Db 280 RLTIPESSCPSPFAELHQCMEADAKKRPSPFOIISIL 316
```

Search completed: December 10, 2002, 03:47:53
Job time : 89 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:44:40 ; Search time 53 Seconds
(without alignments)
549.599 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615
Sequence: 1 MSTASAASSSSSSSAGEMIE.....MTHLMRFPFGADEPLQYPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1615	100.0	567	2	JCS957 transforming growth
2	1615	100.0	579	2	JCS955 transforming growth
3	1615	100.0	606	2	JCS956 transforming growth
4	482.5	29.9	954	1	S68178 mixed-lineage prot
5	434.5	26.9	407	1	G84635 probable protein k
6	434	26.9	412	2	T10671 protein kinase hom
7	430	26.6	1338	2	T18287 protein-tyrosine k
8	427.5	26.5	391	2	T48115 protein kinase ATM
9	421.5	26.1	357	2	C84856 probable protein k
10	419	25.9	847	1	A53800 mixed-lineage prot
11	418.5	25.9	394	2	JU00229 mixed-lineage prot
12	416	25.8	328	2	T16747 hypothetical prote
13	411.5	25.5	546	2	D84555 probable protein k
14	411.5	25.5	553	2	T04683 hypothetical prote
15	411	25.4	1030	2	P96763 hypothetical prote
16	408.5	25.3	410	2	B35670 protein-tyrosine k
17	406.5	25.2	370	2	T46150 protein kinase ATM
18	401	24.8	668	2	JC2163 protein kinase (EC
19	399	24.7	888	2	A55318 serine/threonine p
20	399	24.6	821	2	JCS399 dual leucine zippe
21	398	24.6	829	2	T48400 serine/threonine-p
22	397.5	24.6	829	2	T07406 probable protein k
23	391.5	24.2	736	2	T05137 protein kinase hom
24	391	24.2	364	2	G71410 probable protein k
25	387	24.0	475	2	T12955 probable protein k
26	387	24.0	963	2	T09911 probable serine/th
27	385	23.8	438	2	C86273 protein kinases ho
28	384	23.8	387	2	T22511 hypothetical prote
29	383	23.7	982	2	T06576 probable protein k

30	378.5	23.4	356	2	S61766 protein kinase ATN
31	376.5	23.3	1584	2	T18276 protein-tyrosine k
32	374.5	23.2	406	2	T52626 probable mitogen-a
33	374	23.2	1015	2	T00726 probable serine/th
34	373	23.1	462	2	S29851 protein kinase 6 (
35	370	22.9	1029	2	H86179 hypothetical prote
36	366.5	22.7	390	2	T01451 protein kinase hom
37	365.5	22.6	356	2	T148206 protein kinase ATM
38	365.5	22.6	630	1	T01380 protein-tyrosine k
39	365.5	22.6	631	2	I56997 protein-tyrosine k
40	365	22.6	738	2	F96701 hypothetical prote
41	365	22.6	1171	2	T12956 hypothetical prote
42	363.5	22.5	527	2	S13763 protein-tyrosine k
43	363.5	22.5	1064	1	S57450 protein-tyrosine k
44	362	22.4	1257	2	T00486 serine/threonine-s
45	361	22.4	375	2	D84715 probable protein k

ALIGNMENTS

RESULT 1
JCS957
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C:Accession: JCS957
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243: 545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JCS955; MUID:98153801; PMID:9480845
A:Accession: JCS957
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-567 <SAK>
A:Cross-references: DDBJ:AB009358
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase

Query Match 100.0%; Score 1615; DB 2; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.9e-83;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MSTASAASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEYVGARGVCAKAKRAKV	60
DB	1	MSTASAASSSSSSSAGEMIEAPSOVLNFEEDYKEIEVEYVGARGVCAKAKRAKV	60
QY	61	AIKQIESSEERRAPFVLELRQLSRVNHPNIVKL YGACLNVPCLVMEYAGGSLYNVLHGAE	120
DB	61	AIKQIESSEERRAPFVLELRQLSRVNHPNIVKL YGACLNVPCLVMEYAGGSLYNVLHGAE	120
QY	121	PLPYTTAAHAWMCVQCQSGVAYLHSMOPKAL IHRDLKPPMLLVAGTGVKICDPFGAC	180
DB	121	PLPYTTAAHAWMCVQCQSGVAYLHSMOPKAL IHRDLKPPMLLVAGTGVKICDPFGAC	180
QY	181	DIQTMNTNKKSGAAMMADEVEFGSNYSEKCVFSWGIIIMEVITRRKPFDEIGSPARIM	240
DB	181	DIQTMNTNKKSGAAMMADEVEFGSNYSEKCVFSWGIIIMEVITRRKPFDEIGSPARIM	240
QY	241	WAVHNGTRPPLIKNIPKPIESIMTRCWSKDPSPRSMEEIYKIMTHLMRYPPGADEPLQY	300
DB	241	WAVHNGTRPPLIKNIPKPIESIMTRCWSKDPSPRSMEEIYKIMTHLMRYPPGADEPLQY	300
QY	301	PCQ 303	
DB	301	PCQ 303	

RESULT 2
JCS955
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000

C/Accession: JCS955
R; Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A; Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A; Reference number: JCS955; MUID:98153801; PMID:9480845
A/Accession: JCS955
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-579 <SAK>
A/Cross-references: DBJ:AB009356; NID:g2924623; PIDN:BAA25025.1; PID:g2924624
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: phosphotransferase

Query Match 100.0%; Score 1615; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.9e-83;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKOIESESRKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKOIESESRKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAVLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAVLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIOTHTNNTNKGSAAMWAPEVFGSNYSKCDVPSWGIILWEVITTRKPFDEIGGPAFRIM 240
DB 181 DIOTHTNNTNKGSAAMWAPEVFGSNYSKCDVPSWGIILWEVITTRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
DB 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
JCS956
Transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C/Species: Homo sapiens (man)
C/Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C/Accession: JCS956
R; Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A; Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A; Reference number: JCS955; MUID:98153801; PMID:9480845
A/Accession: JCS956
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-606 <SAK>
A/Cross-references: DBJ:AB009357; NID:g2924625; PIDN:BAA25026.1; PID:g2924626
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C/Keywords: phosphotransferase

Query Match 100.0%; Score 1615; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 2e-83;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
QY 61 AIKOIESESRKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKOIESESRKAFIVELRQLSRVNHPNIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYTTAAHAMSWCLQCSQGVAVLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180

DB 121 PLPYTTAAHAMSWCLQCSQGVAVLHSMQPKALIHRLDKPPNLLLVAGGTVLKICDFGTAC 180
QY 181 DIOTHTNNTNKGSAAMWAPEVFGSNYSKCDVPSWGIILWEVITTRKPFDEIGGPAFRIM 240
DB 181 DIOTHTNNTNKGSAAMWAPEVFGSNYSKCDVPSWGIILWEVITTRKPFDEIGGPAFRIM 240
QY 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
DB 241 WAVHNGTRPPLIKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFPGADPQY 300
QY 301 PCQ 303
DB 301 PCQ 303

RESULT 4
S68178
mixed-lineage protein kinase 2 (EC 2.7.1.-) - human
C/Species: Homo sapiens (man)
C/Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
C/Accession: S68178; I38044; S32468
R; Dorow, D.S.; Devereux, L.; Tu, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simpe
Eur. J. Biochem. 234, 492-500, 1995
A; Title: Complete nucleotide sequence, expression, and chromosomal localisation of humar
A; Reference number: S68178; MUID:96128179; PMID:8536694
A/Accession: S68178
A/Molecule type: mRNA
A/Residues: 1-954 <DOR>
A/Cross-references: EMBL:X90846; NID:g971419; PIDN:CAA62351.1; PID:g971420
R; Katoh, M.; Hirai, M.; Sugimura, T.; Terada, M.
Oncogene 10, 1447-1451, 1995
A; Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinase
A; Reference number: I38044; MUID:95249256; PMID:7731697
A/Accession: I38044
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-461, 'A', 'V', 465-470, 'S', 472-806, 'R', 808-817, 'A', 819-954 <RES>
A/Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA85531.1; PID:g758593
R; Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A; Title: Identification of a new family of human epithelial protein kinases containing t
A; Reference number: S32467; MUID:93238756; PMID:8477742
A/Accession: S32468
A/Molecule type: mRNA
A/Residues: 244-464, 'AQAAGRRQHPALWL' <DO2>
C/Genetics:
A; Gene: GDB:MLK2; GDB:MST
A; Cross-references: GDB:362654; GDB:624810; OMIM:600137
A; Map position: 19q13.1-19q13.2
C/Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F; 23-76/Domain: SH3 homology <SH3>
F; 96-364/Domain: protein kinase homology <KIN>
F; 104-112/Region: protein kinase ATP-binding motif
F; 384-405/Region: leucine zipper motif
F; 419-440/Region: leucine zipper motif
F; 449-463/Region: basic
F; 125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 29.9%; Score 482.5; DB 1; Length 954;
Best Local Similarity 38.4%; Pred. No. 5.3e-20;
Matches 107; Conservative 51; Mismatches 98; Indels 23; Gaps 7;

QY 21 APSQVLNFEIDYKEIEVEEVVGRGAFGVVCKAKWRAKDVVAIKQIESESRKAFIV---- 76
DB 83 APAGLQLPQETPPPELQLEBIIGVGSGKVVYRALWRGEEVAVKAARLDPEKDPAVTAEQV 142
QY 77 --ELRQLSRVNHNPVILKLYGACLNPP--VCLVMEYAEAGGSLYNVLHGAEPLPYTTAAHAMS 132
DB 143 CQEARLFGALQHPNIIALRGACLNPPHCLVMEYARGGALSRLAGRRVPPHV----LVN 198
QY 133 WCLQCSQGVAVLHSMQPKALIHRLDKPPNLLLV-----AGGTVLKICDFGTACD-IQT 184

Db	199	WAVQVARGANNVLLHNDAPFPIITHRDLSINILITLEAIENHNHNLADTVLKTTDFGLAREWHKT	258
Qy	185	HMTNNKSGAAMMADEVEPEGNSYSEKCDVFSWGIIIMLEYITRRKPDFEIGCAPFRIMAV-	243
Db	259	TKMSAAGTYAMMADEVIRSLFSKSGDVMSFGVLLMELLTGEVYREI--DALAAVYQVA	316
Qy	244	HNGTRPFLIKVLPKPIESLMTRCWKSQSPRSMSEIYK	282
Db	317	MNKLTPIPSTCPSEFARLLBECWDPDHPGRPDGCSIIK	355

RESULT 5
G84635
probable protein kinase [imported] - Arabidopsis thaliana
C|Species: Arabidopsis thaliana (mouse-ear cress)
C|Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C|Accession: G84635
R|Lin. X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shenn, M.; Vankar, S.E.; Unayam, L.; Tallon, L.;
Euss, D.; Niernann, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A|Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
#Reference number: A84420; MUID:20083487; PMID:10617197

Query Match	26.9%;	Score 434.5;	DB 2;	Length 407;
Best Local Similarity	34.8%;	Pred. No. 1.2e-17;		
Matches 110; Conservative	54;	Mismatches 97;	Indels 55;	Gaps

RESULT 6
T10671
protein kinase homolog EE21.90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C:Accession: T10671
R:Bayan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, June 1999
A:Reference number: Z16533
A:Accession: T10671

A;Molecule type: DNA
A;Residues: 1-412 <REV>
A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
A;Experimental source: cultivar Columbia; BAC clone F6E21
C;Genetics:
A;Gene: ATSP:F6E21.90
A;Map position: 4
A;Introns: 300/2
C;Superfamily: Kinase-related transforming protein; protein kinase homology
<KIN>
</135-392>Domain:protein kinase homology <KIN>

QY 187 TNNKGSAMMAPEVFEBSNSNSECQDVFSWIIIMEVITRKKFDEIG--PAPRIIMAVH 244
 Db 289 TPERGTYRMMAPEMIQRPYTKVDVYSPGIVLMELITGLPLPQNMNTAVQAAPAV--VN 345
 QY 245 NGTRPPLIKLKPRIEBSLMTRCMSKDPSSQSPSMEELVYKIM 284
 Db 346 RGVRRTPVADCLPVLGEIMTRCWDADPEVAPCPAAELVNL 385

Query Match	26.6%;	Score 430;	DB 2;	Length 1338;
Best Local Similarity	37.7%;	Pred. No. 6e-17;		
Matches 98;	Conservative 50;	Mismatches 94;	Indels 18;	Gaps 7

QY 262 LMTRCWSKDPQRSQPSMEEIV 281
Db 1279 LITKCWASDFSQPSFTEIL 1298

RESULT 8
T48115
protein kinase ATMRL1 (EC 2.7.1.1-) [imported] - Arabidopsis thaliana
N:Alternate names: protein Fl6M2.110
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2001
C:Accession: T48115; T51942
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24459
A:Accession: T48115
A:Molecule type: DNA
A:Residues: 1-391 <RIE>
A:Cross-references: EMBL:AL138648; NID:g7523397; PIDN:CAB86427.1; PID:g7523408
A:Experimental source: cultivar Columbia; BAC clone Fl6M2
R:Ichimura, K.; Mizoguchi, T.; Shinozaki, K.
Plant Sci. 130, 171-179, 1997
A:Title: ATMRL1, an Arabidopsis protein kinase related to mammal mixed-lineage kinases a
A:Reference number: Z24427
A:Accession: T51942
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-391 <ICH>
A:Cross-references: EMBL:AB006810; NID:g2351096; PIDN:BAA22079.1; PID:g2351097
A:Experimental source: cultivar Columbia
C:Genetics:
A:Map position: 3
A:Introns: 109/1; 149/3; 220/2; 278/3; 323/3
A:Note: Fl6M2.110
C:Superfamily: Kinase-related transforming protein; protein kinase homology
C:Keywords: phosphotransferase

Query Match 26.5%; Score 427.5; DB 2; Length 391;
Best Local Similarity 31.3%; Pred. No. 2.9e-17;
Matches 103; Conservative 58; Mismatches 101; Indels 67; Gaps 10;

QY 2 STASASSSSSAGEMIEAPSOVLNPFEEIDYKEIEVEEVGSGAGVYCKAKWRAKDYA 61
Db 55 STSVLNTSASTS-----APAMQEWELDSLKDMDKVLHAGTGTGTVRGVYAGQEVA 108

QY 62 IKQI-----ESESERKAFIVELRQLSRVNHNPVIVKLYGACL----- 97
Db 109 VKVLDMGEDGYATPAETALRASFEQEVAVWQKLDHPNVTKFIGASMGTSDLRIPPAGDT 168

QY 98 -----NPVCLVMEYAEGLSLVNLHGAEPLPYTA-----AHAMSWCLQCSQGVAY 143
Db 169 GGRGNGAHPARACCVVEYVAGGTLKKFL-----IKKYRAKLPIDKVIQLALDLARGLSY 223

QY 144 LHSMQPKALIHRLKPNLLVAGGTVLKICDFGTA---CDIQTHMTNNKSGAAWMAPEV 200
Db 224 LRS---KAIVHRDVKENMLLQPNKT-LKIADFGVARVAQNPDMTGTGTGLGIMAPEV 279

QY 201 FEGSNYSEKCDVFSWGIILWEVITRRKP-----FDEIGGPAPRIMWAVHNGTRPPLIKNL 255
Db 280 LEGKPNRKCDVYSFGVCLWEIYCCDMPYADCQSFABISHA-----VVHRNLRPEIPKCC 333

QY 256 PKFIESLMTRCWSKDPQRSQPSMEEIVKIM 284
Db 334 PHAVANIMKRCWPNDRRPEMEVVKLL 362

RESULT 9
C84856
probable protein kinase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: C84856

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: C84856
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-357 <STO>
A:Cross-references: GB:AE002093; NID:g4559329; PIDN:AAD22991.1; GSPDB:GN00139
C:Genetics:
A:Gene: At2g42630
A:Map position: 2

Query Match 26.1%; Score 421.5; DB 2; Length 357;
Best Local Similarity 36.3%; Pred. No. 5.9e-17;
Matches 106; Conservative 50; Mismatches 85; Indels 51; Gaps 15;

QY 17 BMEIAP----SOVLNPFEE--IDYKEIEVEEVGSGAGVYCKAKWRAKDYAIK-----QI 65
Db 79 EVLESMPQNKPLLPFFBEWNIDFSKLKVGASVSGSGTSGVWVCRGVYNNKTEVAIKIFLQQQL 138

QY 66 ESESERKAFIVELRQLSRVNHNPVIVKLYGACLNP--VCLVMEYAEGLSLVNLHGAEPLP 123
Db 139 TAEN-MKVFCNEISILSRLOHPNVILLGACTKPPQLSLVTEYMWSTGSLDVR----- 191

QY 124 YTTAAHMSW-----CLQCSQGVAYLHSMQPKALIHRLKPNLLVAGGTVLKICDFG 177
Db 192 --TRKELSMQKLLAEICRGLMYTHKM---GIVHRDLTSANCLL--NKSIVKICDFG 244

QY 178 TACDIQTHMTNN-----KGSAAWMAPEVEGSGNYSEKCDVFSWGIILWEVITRRKPPD 230
Db 245 ----LSRRMTGTAVKDTAAAGTPEWMAPELRNPEVPTKSDIFSFGVIMMELSTLSKPKW 300

QY 231 BIGGPAPRIMWAVHN-GTRPPLIKNLPKPIESLMTRCWSKDPQRSQPSMEEIV 281
Db 301 --GVPEKVIHIVANEGAR---LKIPEGPLQKLIADQWS-EPEQRPSCKEIL 346

RESULT 10
A53800
mixed-lineage protein kinase (EC 2.7.1.1-) 3 - human
N:Alternate names: protein kinase PTK1; protein kinase SPRK
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A53800; I58395
R:Gallo, K.A.; Mark, M.R.; Scadden, D.T.; Wang, Z.; Gu, Q.; Godowski, P.J.
J. Biol. Chem. 269, 15092-15100, 1994
A:Title: Identification and characterization of SPRK, a novel src-homology 3 domain-cont
A:Reference number: A53800; MUID:94253068; PMID:8195146
A:Accession: A53800
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-847 <GAL>
A:Cross-references: GB:U07747; NID:g464027; PIDN:AAA19647.1; PID:g464028
R:Ring, Y.L.; Leung, I.W.; Heng, H.H.; Tsui, L.C.; Lassar, N.J.
Oncogene 9, 1745-1750, 1994
A:Title: MLK-3: identification of a widely-expressed protein kinase bearing an SH3 domai
A:Reference number: I58395; MUID:94239754; PMID:8183572
A:Accession: I58395
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-847 <RES>
A:Cross-references: GB:L32976; NID:g488295; PIDN:AAA59859.1; PID:g488296
C:Genetics:
A:Gene: GDB:MLK3; PTK1; SPRK
A:Cross-references: GDB:134755; OMIM:600050
A:Map position: liq3.1-11q3.3
C:Superfamily: mixed-lineage protein kinase 3; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F:48-100/Domain: SH3 homology <SH32>
F:115-383/Domain: protein kinase homology <KIN>

F:123-131/Region: protein kinase ATP-binding motif
F:403-424/Region: leucine zipper motif
F:438-459/Region: leucine zipper motif
F:468-482/Region: basic

Query Match 25.9%; Score 419; DB 1; Length 847;
Best Local Similarity 34.6%; Pred. No. 1.7e-16;
Matches 102; Conservative 53; Mismatches 102; Indels 38; Gaps 9;

QY 22 PSQVLFNEEDYKIEVEEVGAGFVCAKWRADVAIK-----QISESESR 71
Db 108 PCVASFQ-----ELRLEEVIGIGFGKYYRSGMRGELVAVKAADPPEDISVTAESIR 162
QY 72 KAFIVLRQLSRVNHPIYKLYGACINP--VCLVMEYAGSSLYNVLHGAEPLPYTAAH 129
Db 163 Q-----EARLFAMLAHPNITALLKAVCLSEBPNLCVMEYAGGPLSRALAGRVRPHV---- 214
QY 130 AMSWCLQSGGVAYLHSMQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD- 181
Db 215 LVNNAVQIRGMHYHLCALVPVIRHDKSNLILLQPIESDMEHKTKITDFTGLAREW 274
QY 182 IOTHTNNGSAAMAPEVEGSSNYSKCDVFSWGIIMEVITRRKPFDEIGGPAFRIMW 241
Db 275 HKTTQMSAAGTAVAMAPETIKASTSKSDVMSFGVLMELLTGEVPRGIDCLAVANGV 334
QY 242 AVHNGTRPPLIKNLKPIESLMTRCMSKDPGSPS---MEEIVKIMTHLMKYPF 232
Db 335 AVNKLTL-PIPTCEPFAQLMADCAODPHRRDPFASITLQLEALAEQVLRMP 368

RESULT 11

JU0229
mixed-lineage protein kinase 1 - human
C/Species: Homo sapiens (man)
C/Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
C/Accession: S32467; JU0229
R/Dowry, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213: 701-710, 1993
A>Title: Identification of a new family of human epithelial protein kinases containing
A/Reference number: S32467; MUID:93238756; PMID:8477742
A/Accession: S32467
A/Molecule type: mRNA
A/Residues: 1-394 <DO2>
C/Genetics:
A:Gene: GDB:MLK1
A/Cross-references: GDB:141921; OMIM:600136
A/Map position: 14q24.3-14q31
C/Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C/Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F:1-269/Domain: protein kinase homology <KIN>
F:1-268/Domain: catalytic <CAT>
F:9-17/Region: protein kinase ATP-binding motif
F:289-310/Region: leucine zipper motif
F:324-345/Region: leucine zipper motif
F:354-368/Region: basic

Query Match 25.9%; Score 418.5; DB 2; Length 394;
Best Local Similarity 36.6%; Pred. No. 9.4e-17;
Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;

QY 35 EIEVEEVGRAPGVCAKWRADVAIK--QISESESRKAFIYELN-----LSRVNHPN 88
Db 2 ELTLEIRIGIGFGKYYAFAPIGDEVAVKARHPDEDISGTINNVQEAFLFMLKHPN 61
QY 89 IVKLYGACINP--VCLVMEYAGSSLYNVLHGAEPLPYTAAHMSCLQSGGVAYIHS 146
Db 62 IIALRGVCLKEPNLCVMEFARGPLNRVLSGKRIPDI---LVNNAVQIARGMNVLHD 117
QY 147 MQPKALIHRLDKPNNLLV-----AGTVLKICDFTGACD-IQTHMTNNGSAAMAP 198
Db 118 EATVPILHRDKSNLILLQVENGDISNKLITDFTGLAREMHTTRKMSAAGTYAMMAP 177
QY 199 EVFEGSNYSKCDVFSWGIIMEVITRRKPFDEIGGPAFRIMAV--HNGTRPPLIKNLK 257

Db 178 EVIAMSWSKSGDVMSYGVLLMELLTGEVFRGIDG--LRAVAGVAMNKLALPIPTCP 235
QY 258 PIESLMTRCMSKDPGSPSMEIYKIMT 285
Db 236 PFAKLMECWNPDPHSRPSFTNILDQLT 263

RESULT 12

116747
hypothetical protein R13F6.7 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 17-Mar-2000
C/Accession: T16747
R/Miller, N.
submitted to the EMBL Data Library, April 1994
A/Description: The sequence of C. elegans cosmid R13F6.
A/Reference number: Z18570
A/Accession: T16747
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1328 <ML>
A/Cross-references: EMBL:U00046; NID:9470358; PID:9470364; PIDN:AA047047.1; GSPDB:GN0002
A/Experimental source: strain Bristol N2; clone R13F6
C/Genetics:
A:Gene: CESP:R13F6.7
A/Map position: 3
A/Introns: 20/3; 160/3; 222/2; 286/2
C/Superfamily: protein-tyrosine kinase src; protein kinase homology; SH2 homology; SH3 h

Query Match 25.8%; Score 416; DB 2; Length 328;
Best Local Similarity 33.4%; Pred. No. 1.1e-16;
Matches 99; Conservative 57; Mismatches 112; Indels 28; Gaps 11;

QY 1 MSTRASASSSSSSAGBMTAPSOVLNFEEDYKIEVEEVGAGFVCAKWRADV 60
Db 22 MSTPTSNESSTSSSNNNS-----DQVLPFDIQDDIQVDHIGVGFAGVFGSNWTLPGD 76
QY 61 AIKQISESESRKAFIYELN--LSRVNHPNIVKLYGACN--NPVCLVMEYAGSSLYNV 115
Db 77 SQRTI--ALKKVFLEKEAEILSKIRHNIIQFYGICATGNDPFIYVEAKESLYDF 133
QY 116 LHGAEPPLPYTAAHMS-----WCLQSGGVAYLHSMQPKALIHRLDKPNNLLVAGGT 169
Db 134 IHSSESGSFASSSGNSPVPVYKMAQIASGYQLHYDAVDITIIHRDLKSKVVL-DKVL 192
QY 170 VLKICPFTACDIQTHMTNNK--GSAAMAPD-VFEGSNYSKCDVFSWGIIMEVITR 225
Db 193 VCKICDFGSKDL-THSCTPASWGTAAWMSPEMLIQSEGLTTATDVMSYGVLLWEIISK 251
QY 226 RKPFDEIGGPAFRIMAV--HNGTRPPLIKNLKPIESLMTRCMSKDPGSPSMEI 280
Db 252 EVYKYKS--ERRIFTMTQSGITLAIIPSCPAPLKQUMSNCMKMTPKDRANMROI 305

RESULT 13

D84555
probable protein kinase (imported) - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C/Accession: D84555
R/Jin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vankken, S.E.; Unayam, L.; Tallon, L.
eaus, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402: 761-768, 1999
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: D84555
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-546 <STO>
A/Cross-references: GB:AE002093; NID:96598802; PIDN:AA080785.2; GSPDB:GN00139
C/Genetics:

A:Gene: At2g17700
A:Map position: 2

Query Match 25.5%; Score 411.5; DB 2; Length 546;
Best Local Similarity 33.3%; Pred. No. 3e-16;
Matches 96; Conservative 61; Mismatches 116; Indels 15; Gaps 8;
QY 10 SSSSAGEMIEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDVAIKQIES 69
DB 260 SNEILPACIEIPTDGTDEWEIDVQLKTEKKVAGSYGDLHRTGYCQEVAKFLKPD 319
QY 70 ERKAFIVELRQ----LSRVNHNPNIVKLYGACLN--PVLWMEYAEAGSLYNVLHCAEPLP 123
DB 320 VNNEMLREFSQEVFMKVRHKNVQFLGACTRSPCLTCLVTFEMARGSIYDFLHKQKCA- 378
QY 124 YTTAAHAMSWCLQCSOGVAYLHSMQPKALIHRLDKPPLNLLVAGGTGVLKICDFGTA-CDI 182
DB 379 -FKLQTLKVALDVAKGMSYLHQ---NNIIHRDLKTANLLMDEHGLV-KVADFGVARVQI 433
QY 183 QTH-MTNNKGSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMW 241
DB 434 ESOVMTAETGYRWMAPEVIEHKPYNHKADVFSYAIVLWELLTGDIPIYAFV-TPLQAAVG 492
QY 242 AVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSPQSMEEIVKIMTHLMR 289
DB 493 VVOKGLRPKIPKTKTHPKVKGLLERCHQDPQEPQRFEEIEMLQOIMK 540

RESULT 14

T04683
hypothetical protein F8D20.290 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-May-1999
C:Accession: T04683
R:Bevan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Jesse, T.; Heijnen, L.; Vos, P.; Mewes, submitted to the Protein Sequence Database, July 1998
A:Reference number: Z15381
A:Accession: T04683
A:Molecule type: DNA
A:Residues: 1-553 <BEV>
A:Cross-references: EMBL:AL031135
A:Experimental source: cultivar Columbia; BAC clone F8D20
C:Genetics:
A:Map position: 4
A:Introns: 69/2; 107/3; 176/2; 194/3; 231/3; 289/2; 325/2; 350/1; 408/3; 467/3; 4
A:Note: F8D20.290

Query Match 25.5%; Score 411.5; DB 2; Length 553;
Best Local Similarity 32.2%; Pred. No. 3.1e-16;
Matches 96; Conservative 60; Mismatches 111; Indels 31; Gaps 9;
QY 19 IEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDVAIKQIESESEKAFIVEL 78
DB 258 VEIPTDGTDEWEIDVQLKTEKKVAGSYGELFRGTGYCQEVAKILKPERVNAEMLREF 317
QY 79 RQ----LSRVNHNPNIVKLYGACLN--PVLWMEYAEAGSLYNVLHCAEPLPYTTAAHMS 132
DB 318 SQEYIMRKVRHKNVQFLGACTRSPCLTCLVTFEMRGSYDFLHKHKGV--FKIOSLLK 375
QY 133 WCLQCSOGVAYLHSMQPKALIHRLDKPPLNLLVAGGTGVLKICDFGTACDIQTH---MTNN 189
DB 376 VALDVSKGMNYLHQ---NNIIHRDLKTANLLM-DEHVVVKVADFGVA-RVQTESGVMTAE 430
QY 190 KGSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWVHNGTRP 249
DB 431 TGTWRMAPEVIEHKPYNDHRAVFSYAIVLWELLTGELPYSYL-TPLQAAVGVVQKGLRP 489
QY 250 PLIKNLPKPIESLMTRCWSKDPQSPQSMEEIVKIMTHLMR-----YFPG 293
DB 490 KIPKETHPKLTLELCKECQDPALRPFAETIEMNLQIREVIDLSLHKDKHGGYFSG 547

RESULT 15

F96763

hypothetical protein F25P22.8 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96763
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Dewar, K.; Jensen, N.P.; Hughes, B.; Huijzer, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ket, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: AB6141; MUID:21016719; PMID:11130712
A:Accession: F96763
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1030 <STO>
A:Cross-references: GB:AE005173; NID:G6692730; PIDN:AAF24836.1; GSPDB:GN00141
C:Genetics:
A:Gene: F25P22.8
A:Map position: 1

Query Match 25.4%; Score 411; DB 2; Length 1030;
Best Local Similarity 32.9%; Pred. No. 5.5e-16;
Matches 103; Conservative 57; Mismatches 125; Indels 28; Gaps 10;
QY 1 MSTASASSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVVGAFGVCKAKWRAKDV 60
DB 721 ISDKSIGNSSKSDCDDVSDC-----ELLWEITVGERIGLSYGEYVRGDMHGTEV 772
QY 61 AIKQIESE---SERKAFIVELQLSRVNHNPVTKLYGACLN--VCLWMEYAEAGSLYN 114
DB 773 AVKFLDQDLTGEALBEFRSEVRIMKLRHPNIVLFMGAVTRPPNLSIVTEFLRGSLSYR 832
QY 115 VLHGAELPLPYTTAAHMSWCLQCSOGVAYLHSMQPKALIHRLDKPPLNLLVAGGTGVLKIC 174
DB 833 LIH--RPNQLDERRRLMALDARGNNYLHSCNP-MIVHRDLKSPN-LLVDKNWVVKVC 888
QY 175 DFGTACDIQTHMT-----NNKGSAAWMAPEVFECSNYSEKCDVFSWGIILWEVITRRKPF 229
DB 889 DFGLS--RMKHSTVLSKSTAGTAGTAENWMAPEVLRNEPADEKCDVSYGVILWELFTLQQPW 946
QY 230 DEIGGPAFRIMWVHNGTRPPLIKNLPKP-LESIMTECWSKDPQSPQSMEEIVKIMTHLM 288
DB 947 GKMN--PMQVVGVAGVGFQHRRLDIPDFVDPALADLSKQWOTDSKLRPSFAEIMASLKRQ 1004
QY 289 RYFFGADEPLQYP 301
DB 1005 KPTVGSNIPRPVP 1017

Search completed: December 10, 2002, 03:51:45
Job time : 55 secs

KW Alternative splicing.
FT DOMAIN 8 14
FT NP_BIND 36 291
FT BINDING 42 50
FT BINDING 63 63
FT ACT_SITE 156 156
FT VARSPLIC 404 430
FT VARSPLIC 509 518
FT VARSPLIC 519 606
SQ SEQUENCE 606 AA; 67196 MW; 308F8147CD174013 CRC64;

Query Match 100.0%; Score 1615; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 7e-134;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60

QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120

QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGVTLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGVTLKICDFGTAC 180

QY 181 DIQTHMTNNKGSAAWMAPEVFEESNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEESNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWSDQSPQSPMSBEEIVKIMTHLMRYFFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWSDQSPQSPMSBEEIVKIMTHLMRYFFPGADEPLQY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 2
M3K7_MOUSE STANDARD; PRT; 579 AA.
ID M3K7_MOUSE
AC Q62073;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-activated kinase 1).
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=96123277; PubMed=8533096;
RX Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N., Taniuchi T., Nishida E., Matsumoto K.;
RT Identification of a member of the MAPKK family as a potential mediator of TGF-beta signal transduction.";
RL Science 270:2008-2011(1995).
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC -!- MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B ACTIVATION.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; D76446; BAAL1184.1; .
CC HSSP; P08631; LADS.
CC GMD; MG1:1346877; Map3k7.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYKC_1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC TRANSFERASE; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 8 16
FT DOMAIN 36 291
FT NP_BIND 42 50
FT BINDING 63 63
FT ACT_SITE 156 156
FT ACT_SITE 156 156
SQ SEQUENCE 579 AA; 64227 MW; 97C8F6F3C8E283EE CRC64;

Query Match 99.6%; Score 1609; DB 1; Length 579;
Best Local Similarity 99.7%; Pred. No. 2.2e-133;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60
DB 1 MSTAASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVVCKAKWRAKDV 60

QY 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120
DB 61 AIKQIESESEKAFIVELRQLSRVNHPIVLYGACLNVPCLVMEYAEAGSLYNVLHGAE 120

QY 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGVTLKICDFGTAC 180
DB 121 PLPYTTAAHAMSWCLQCSQGVAYLHSMQKALIHRLDKPNNLLVAGGVTLKICDFGTAC 180

QY 181 DIQTHMTNNKGSAAWMAPEVFEESNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNNKGSAAWMAPEVFEESNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWSDQSPQSPMSBEEIVKIMTHLMRYFFPGADEPLQY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTKRCWSDQSPQSPMSBEEIVKIMTHLMRYFFPGADEPLQY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
M3K7_MOUSE STANDARD; PRT; 393 AA.
ID M3K7_MOUSE
AC P83104;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative mitogen-activated protein kinase kinase 7 (EC 2.7.1.-).
GN TAK1L
OS Drosophila melanogaster (Fruit fly)
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkeley;
RC MEDLINE=20196006; PubMed=10731132;
RX

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amatiello P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chame M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abriil U.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart M.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck U.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclev J.M.,
 RA Palazzolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Manning G., Sudarsanam S., Ploeman G.;
 RT "Prediction of novel protein kinases from the *Drosophila* genome
 RL project and EST sequences.";
 RT Unpublished observations (Aug-2001).
 CC -!- FUNCTION: Can phosphorylate and activate yet undefined MAPKs.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE003732; -!- NOT ANNOTATED_CDS.
 DR FlyBase: FBgn0046689; Tak1.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR002290; Ser_thr_kinase.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00069; Pkinase.1.
 DR PRINTS: PR00109; TYRKINASE.
 DR PRODOM: PD000001; Euk_kinase.1.
 DR SMART: SM00220; S_TKc; 1.
 DR SMART: SM00219; TYKc; 1.
 DR PROSITE: PS00107; PROTEIN KINASE ATP.1.
 DR PROSITE: PS00101; PROTEIN KINASE DOM.1.
 DR PROSITE: PS00108; PROTEIN KINASE ST.1.
 KW Hypothetical protein; Transferase; Serine/threonine-protein kinase;
 KW Tyrosine-protein kinase; ATP-binding.
 FT DOMAIN 11 266 PROTEIN KINASE.
 NP_BIND 17 25 ATP (BY SIMILARITY).

FT BINDING 38 38 ATP (BY SIMILARITY).
 FT ACT_SITE 133 133 BY SIMILARITY.
 SQ SEQUENCE 393 AA; 45238 MW; 43370C98490F0384 CRC64;
 Query Match 30.1%; Score 486; DB 1; Length 393;
 Best Local Similarity 36.5%; Pred. No. 3.3e-35;
 Matches 105; Conservative 66; Mismatches 91; Indels 24; Gaps 9;
 QY 29 EIDYKEIEV-EEVVGKAPGVCKARKDAIK-----QIESESEKATIVELRQ 80
 DB 3 KQVDAEYKLESEKFLGSGGAVRKATFQNEIAVKLFDFLEETIKKNAE-----EITH 57
 QY 81 LSRVNHPIVLKACLN--PVCIMVEAEGSSYVNLHGAEPPLPYTAAHAWMCQCS 138
 DB 58 LSEIDHEVIVIGVIGASGKKDYILMELEBGSJANLYGDKWE--TVDEAVRWALQCA 116
 QY 139 QGVAVLHSMOPKALIHRLKRPENLLVAGGVTLICPGFACDIDTMTNKSAAWMAP 198
 DB 117 KALATLHSLD-RPIVHRDKIQNMILLVQHEDLKICDFGLATMSNKKTDMDGLRTMAP 175
 QY 199 EYFEGSNSEKCDVFSWGIILMEVITRRKPEDEIGP--AFRIWAVHNGTRPPL---IK 253
 DB 176 EAIKHLKYTANCDVYSFGIMMELMTRLQPSHLNENPSQVAKKAISSGKELPMEAVRS 235
 QY 254 NLPKPIESLMTRCWSKDSQSPSMEIYKINTLMRYFPGADEPLQIP 301
 DB 236 DCEBIGIKOLMECCMDINPEKRPMSKEIEKFLGE--QYEGSTDEDFIXP 281
 RESULT 4
 ID M3XA_HUMAN STANDARD; PRT; 954 AA.
 AC Q02779; Q12761; Q14871;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Mitogen-activated protein kinase kinase kinase 10 (EC 2.7.1.37)
 DE (Mixed lineage kinase 2) (protein kinase MST).
 GN MAP3K10 OR MLK2 OR MST.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 ON NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96128179; PubMed=8536694;
 RA Doriow D.S., Devereux L., Tu G.F., Price G., Nicholl J.K.,
 RA Sutherland G.R., Simpson R.J.;
 RT "Complete nucleotide sequence, expression, and chromosomal
 RL localisation of human mixed-lineage kinase 2.";
 RL Eur. J. Biochem. 234:492-500(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=95249256; PubMed=7731697;
 RA Katoh M., Hirai M., Sugimura T., Terada M.;
 RT "Cloning and characterization of MST, a novel (putative)
 RL serine/threonine kinase with SH3 domain.";
 RL Oncogene 10:1447-1451(1995).
 RN [3]
 RP SEQUENCE OF 244-480 FROM N.A.
 RC TISSUE=colon epithelium;
 RX MEDLINE=93238756; PubMed=8477742;
 RA Doriow D.S., Devereux L., Dietzsch E., de Kreiser T.;
 RT "Identification of a new family of human epithelial protein kinases
 RT containing two leucine/isoleucine-zipper domains.";
 RL Eur. J. Biochem. 213:701-710(1993).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAIN AND SKELETAL MUSCLE.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE KINASE SUBFAMILY.
 CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

```
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X90846; CAA62351.1; --
CC EMBL; Z48615; CAA88531.1; --
CC PIR; S32468; S32468.
CC HSSP; P11362; 1FGK.
CC Genew; HGNC:6849; MAP3K10.
CC MIM; 600137; --
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001452; SH3.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00018; SH3_1.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00452; SH3DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Euk_pkinase; 1.
CC ProDom; PD000066; SH3_1.
CC SMART; SM00326; SH3_1.
CC SMART; SM00326; SH3_1.
CC SMART; SM00221; STYKC_1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS50002; SH3_1.
CC Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
CC ATP-binding; SH3 domain.
CC FT DOMAIN 2 5 POLY-GLU.
CC FT DOMAIN 16 81 SH3.
CC FT DOMAIN 98 360 PROTEIN KINASE.
CC FT NP_BIND 104 112 ATP (BY SIMILARITY).
CC FT BINDING 125 125 ATP (BY SIMILARITY).
CC FT ACT_SITE 222 222 BY SIMILARITY.
CC FT DOMAIN 384 405 LEUCINE-ZIPPER 1 (BY SIMILARITY).
CC FT DOMAIN 419 440 LEUCINE-ZIPPER 2 (BY SIMILARITY).
CC FT DOMAIN 449 463 ARG/LYS-RICH (BASIC).
CC FT CONFLICT 462 464 SRL -> AV (IN REF. 2).
CC FT CONFLICT 465 480 LKREGGSHSLPSGF -> AQAGRRQHPQALWL (IN
CC REF. 3).
CC FT CONFLICT 471 471 G -> S (IN REF. 2).
CC FT CONFLICT 807 807 G -> R (IN REF. 2).
CC FT CONFLICT 818 818 V -> A (IN REF. 2).
CC SEQUENCE 954 AA; 103623 MW; 538F4AAA559B0ABA CRC64;
CC
CC Query Match 29.9%; Score 482.5; DB 1; Length 954;
CC Best Local Similarity 38.4%; Pred. No. 1.9e-34;
CC Matches 107; Conservative 51; Mismatches 98; Indels 23; Gaps 7;
CC
CC QY 21 APSQVLNFEIDYKEVEEVRGAGFVGVCKAKRAKDAIKVAIKQIESESEKAFIV--- 76
CC DB 83 APAGLQLPQIIPHEQLQLEIIGVGFGKVRALWGEVAVAKAARLDPEKPAVTAEOV 142
CC
CC QY 77 --ELRQLSRVNHNPVVKLYGACLNLP--VCLVMEYAGGSLYVNLHGAEPPLPYTTAAHAMS 132
CC DB 143 CQEARLFGALQHNPITALRGACLNPPHCLVMEYARGGALSRLVLAGRRVPPHV---LVN 198
CC
CC QY 133 WCLQCQGVAYLHSMQPKALIHRLDKPPNLLV-----AGGTVLKICDFTGACD-IQT 184
CC DB 199 WAVQVARGMNYLNDAPVPIIHRDLKSNILILQKVENGDLSNLIKITDFGLAREWHRTYGA 258
CC
CC QY 185 HMTNNKGSAAWMAPEVFEFSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV- 243
CC DB 259 TKMSAAGTYAWMAPEVIRLSFSKSDVMSFGVLLNELLITGETVPEYI--DALAVAYGVA 316
CC
CC QY 244 HNTTRPLINKPLKPKIESLMTRCWSKDPQSPMSIEIVK 282
```

```
Db 317 MNKLTLPSTPCPEPFARLLREECWDPDPHGRDFGSILK 355
RESULT 5
ID M3K9 HUMAN STANDARD; PRT; 394 AA.
AC P80192;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 9 (EC 2.7.1.1-) (Mixed
DE lineage kinase 1) (Fragment).
GN MAP3K9 OR MLK1 OR PRKE1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon epithelium;
RX MEDLINE=93238756; PubMed=8477742;
RA Dorow D.S., Devereux L., Dietzsch E., de Kretser T.;
RT "Identification of a new family of human epithelial protein kinases
RT containing two leucine/isoleucine-zipper domains.";
RL Eur. J. Biochem. 213:701-710(1993).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN EPITHELIAL TUMOR CELL LINES OF
CC COLONIC, BREAST AND OESOPHAGEAL ORIGIN.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC PIR; S32467; S32467.
CC PIR; JU0229; JU0229.
CC HSSP; P12931; 1FMK.
CC Genew; HGNC:6861; MAP3K9.
CC MIM; 600136; --
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00219; Tyrc; 1.
CC PROSITE; PS00107; PROTEIN KINASE ATP; 1.
CC PROSITE; PS00108; PROTEIN KINASE ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
CC ATP-binding.
CC NON_TER 1 1
CC DOMAIN 3 271 PROTEIN KINASE.
CC NP_BIND 9 17 ATP (BY SIMILARITY).
CC BINDING 30 30 ATP (BY SIMILARITY).
CC ACT_SITE 127 127 BY SIMILARITY.
CC DOMAIN 289 310 LEUCINE-ZIPPER 1 (BY SIMILARITY).
CC DOMAIN 324 345 LEUCINE-ZIPPER 2 (BY SIMILARITY).
CC DOMAIN 354 368 ARG/LYS-RICH (BASIC).
CC SEQUENCE 394 AA; 44975 MW; DBE40B7D31047FD8 CRC64;
CC
CC Query Match 25.9%; Score 418.5; DB 1; Length 394;
CC Best Local Similarity 36.6%; Pred. No. 2.7e-29;
CC Matches 98; Conservative 47; Mismatches 100; Indels 23; Gaps 8;
CC
CC QY 35 BIEVEEVGRGAGFVGVCKAKRAKDAIK--QIESESEKAFIVELRQ----LSRVNHPN 88
CC DB 2 ELTLEEIIIGVGFGKVRALWGEVAVAKAARHDPDDDISQTIENVRQEAFLFAMLKHPN 61
CC
CC QY 89 IVKLYGACLNLP--VCLVMEYAGGSLYVNLHGAEPPLPYTTAAHAMSCLQCQGVAYLHS 146
CC DB 62 IIALRGVCLKEPNICLVMEFARGGPLNRVLSGKRIPDDI----LVNWAQIARGMNYLHD 117
CC
CC QY 147 MQPKALIHRLDKPPNLLV-----AGGTVLKICDFTGACD-IQTHMTNNKGSAAWMA 198
CC DB 118 EAVIPIHRLDKSNILILQKVENGDLSNLIKITDFGLAREWHRTYGAAGTYAWMA 177
CC
CC QY 199 EVFEFSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMWAV-HNGTRPLINKLPK 257
```

```
Db 178 EVIRASMFSGSDVWSGVILMELLTGEVDFRIGD--LRVAVGANMLALPISTCPE 235
QY 258 PIESLMTRCWSKDSQSPSPSMEELVKIMT 285
Db 236 PFAKLMEDCWNPPHRSRPSFTNILDOLT 263

RESULT 6
KXK2 D1C1D1 STANDARD; PRT; 410 AA.
AC P18161;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYK2 OR DPYK2.
OS Dictyostelium discoideum (Sjime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -! SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33784; AAA33203.1; -.
CC PIR: B35670; B35670.
CC HSSP: P08631; 1AD5.
CC DictyDB: DD03011; pyk2.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR004040; STY_pkinase.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00221; STYK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Transfaser: Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 1 1
FT NON_TER 1
FT DOMAIN 108 381 PROTEIN KINASE.
FT NE_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E93918B6059AACL CRC64;

Query Match 25.3%; Score 408.5; DB 1; Length 410;
Best Local Similarity 33.8%; Pred. No. 2,1e-28;
Matches 94; Conservative 52; Mismatches 103; Indels 29; Gaps 9;

QY 28 FEE-----IDYKEIEVEEVGAFGVYCKAKMAKVAIKOI-----ESSSEK 72
Db 89 FEEKSLTGEREVIITDNDIOPTOKVGEAPSEVSGMWMGHHAIKLIKTIJDDEDFKE 148
QY 73 AFYELRQLSRVNHPIVTKYGACLNPFVCLMEYAEKSGSLYNVLH--GAEPUPYYTAH 129
Db 149 RFIIEVGNLKKGNHNTVMFYGACYPACITTEYMAGSGSLYNILHNNSSTPRVKYSFPL 208
QY 130 AMSWGLQSGVAVLHSMQPKALIHRLDKPPNLLVAGGVTLVKICDFGTACD--IOTHT 187
```

```
Db 209 VLKATDMALGLLHLHET---TIYHRDLTSGNILLDELGNI-KISDGLSAEKREKSM 264
QY 188 NNKG---SAAMMADEVEEG-SNYSEKCDVPSWGIIIMVEVITRRKPEIGGAPRIMAV 243
Db 265 MTNGGICNPRWRPPELFTKNIHGYSEKVDVYCFSLVWEIITLGEIIPFDLDG-SQSAQVA 323
QY 244 HNGTRPPLIKNLPKPIESLMTRCWSKDSQSPSPSMEELV 281
Db 324 YAGLRPPIEYCDPELKLTLQCEWADPNDRPPPTIYV 361

RESULT 7
M3KC_RAT
ID M3KC_RAT STANDARD; PRT; 888 AA.
AC Q63796;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 12 (EC 2.7.1.37)
DE (MAPK-upstream kinase) (MUK).
GN MAP3K12 OR MUK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9626099; PubMed=8637721;
RA Hirai S., Izawa M., Osada S., Spyrou G., Ohno S.;
RT "Activation of the JNK pathway by distantly related protein kinases,
RT MEKK and MUK."
RL Oncogene 12:641-650(1996).
CC -! FUNCTION: May be an activator of the JNK/SAPK pathway.
CC phosphorylates beta-casein, histone 1 and myelin basic protein in
CC vitro.
CC -! CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -! COPACTOR: Magnesium.
CC -! SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated (By
CC similarity).
CC -! PM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC under basal conditions and dephosphorylated when membrane-
CC associated (By similarity).
CC -! SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D49785; BAA08621.1; -.
CC HSSP: P12931; IFMK.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR004040; STY_pkinase.
CC InterPro: IPR002290; Ser_thr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00221; STYK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
CC PROSITE: PS00109; PROTEIN_KINASE_ST; 1.
CC PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
CC Transfaser: Serine/threonine-protein kinase; ATP-binding;
CC Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399 PROTEIN KINASE.
FT NP_BIND 164 172 ATP (BY SIMILARITY).
FT BINDING 185 185 ATP (BY SIMILARITY).
FT ACT_SITE 269 269 BY SIMILARITY.
FT DOMAIN 56 62 POLY-GLY.
FT DOMAIN 668 671 POLY-PRO.
FT DOMAIN 698 701 POLY-PRO.
```


CC	STRAIN=ICR x Swiss Webster; TISSUE=Brain;
RC	MEDLINE=96365388; PubMed=8769565;
RA	Brouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
RT	"Cell-specific expression of the ZPK gene in adult mouse tissues.";
RL	DNA Cell Biol. 15:631-642(1996).
RN	[3]
RP	PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX	MEDLINE=96279269; PubMed=8663324;
RA	Mata M., Merritt S.E., Fan G., Yu G.G., Holzman L.B.;
RT	"Characterization of dual leucine zipper-bearing kinase, a mixed
RT	leucage kinase present in synaptic terminals whose phosphorylation
RT	state is regulated by membrane depolarization via calcineurin.";
RL	J. Biol. Chem. 271:16888-16896(1996).
CC	- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC	Phosphorylates beta-casein, histone I and myelin basic protein in
CC	vivo.
CC	- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC	- COFACTOR: Magnesium.
CC	- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC	- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart,
CC	testis, gastrointestinal tract, stomach, liver and pancreas.
CC	Within the nervous system, predominantly expressed in neurons and
CC	enriched in synaptic terminals.
CC	- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol
CC	under basal conditions and dephosphorylated when membrane-
CC	associated.
CC	- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC	MAP KINASE KINASE SUBFAMILY.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@lsb.sib.ch).
CC	-----
DR	EMBL, U14636; AAA57280.1; -.
DR	EMBL, U23789; AAB17123.1; -.
DR	HSSP, P12931; IFMK.
DR	MGI:1346881; Map3k12.
DR	MGD; MG1:004040; STY_pkinase.
DR	InterPro: IPRO00719; Euk_pkinase.
DR	InterPro: IPRO04040; STY_pkinase.
DR	InterPro: IPRO02290; Ser_thr_pkinase.
DR	Pfam: PF00069; pkinase; 1.
DR	Prodom: PD000001; Euk_pkinase; 1.
DR	SMART: SM00221; STRYC; 1.
DR	PROSITE: PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR	PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
KW	PROSITE: PS50011; PROTEIN KINASE DOM; 1.
KW	Transferrase, Serine/threonine-protein kinase; ATP-binding;
KW	Phosphorylation; Magnesium; Membrane.
FT	DOMAIN 158 399
FT	NP_BIND 164 172
FT	BTNDING 185 185
FT	ACT_SITE 269 269
FT	DOMAIN 56 62
FT	DOMAIN 668 671
FT	DOMAIN 698 701
FT	DOMAIN 753 758
FT	MUTAGEN 185 185
FT	MUTAGEN 192 192
FT	CONFLICT 18 18
FT	CONFLICT 28 29
FT	CONFLICT 382 382
FT	CONFLICT 494 495
FT	CONFLICT 517 517
FT	CONFLICT 794 794
SQ	SEQUENCE 888 AA; 96083 MW; CFCPCFD34F889AAB CRC64;
Query Match	24.8%; Score 401; DB 1; Length 888;
Best Local Similarity	33.5%; Pred. NO.2,4e-27;
Matches 88; Conservative	54; Mismatches 99; Indels 22; Gaps 9;

Oy	30	EIDVVEIEVEEVEVGRGAFVGVCKAKMAKVAIAIQISESERKAFIYELQOLSVHPNI	89
Db	152	EVFPPEILDLQWVSGSGAGVFLGRHFGEBAVAKVNDLKE-----TDIGHRLKRLKHPNI	206
Oy	90	VKLVGACINP--VCLVMEYAGESLYNVLHGAEPLPYRTAAHMSWCLQSGVAYIASHM	147
Db	207	ITFGKVCQTQADPCYCCTILMEFCCQOGLYEVLRAGRPV---TPSLVDWMGMJAGGNNYLIHL	263
Oy	148	QPKALIHHDLLKPPULLLVAGSTVLIKIDCFGRACDIOGHMNNK--GSAAMAPVFPFGSN	205
Db	264	K---IIHEDLKSPP-MLITVDVVKISDFGSKLSKSKTMSFAGVAMAPVIRNEP	319
Oy	206	YSEKDVFSWMIILMEVITRKRPEDEICGPAFRIMAAV-HNGTRPPLIKMLPKPIESLMT	264
Db	320	VSEKDVMSFCVAVMLLETTGEPYKVDVSSA--IIWVGSNLSHLPIVSSCPDGFKLLR	377
Oy	265	RCWSKDPSPQSPSMEEIVKIMTHL	287
Db	378	QOWNSKPRNRPSPFQO---ILHL	397
RESULT 10			
Ctrl	ARATH	STANDARD;	PRT; 821 AA.
ID	Ctrl ARATH	STANDARD;	PRT; 821 AA.
AC	Q05609;		
DT	01-NOV-1995 (Rel. 32, Created)		
DT	01-NOV-1995 (Rel. 32, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Seine/threonine-protein kinase Ctrl1 (EC 2.7.1.37).		
GN	Ctrl1 OR A15G03730 OR F17C15.150.		
OS	Arabidopsis thaliana (Mouse-ear cress).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	eucosids II; Brassicales; Brassicaceae; Arabidopsis.		
CK	NCBI_TaxID=3702;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=cv. Columbia;		
RC	MEDLINE=21016721; PubMed=11130714;		
RX	Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,		
RA	Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,		
RA	Kohara M., Matsumoto M., Matsuno A., Muraki A., Nakayama S.,		
RA	Nakazaki N., Naito K., Okumura S., Shino S., Takenuchi C., Wada T.,		
RA	Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,		
RA	Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,		
RA	Haeblerum K., Murray J., Johnson D., Rohlfing T., Nelson J.,		
RA	Steenken T., Pepin K., Spielhorn J., Sekhon M., Armstrong J., Becker M.,		
RA	Buiter H., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,		
RA	Du H., Edwards J., Fryman J., Heakensen B., Lamar E., Latreille P.,		
RA	Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,		
RA	Wagner-McPherson C., Wollam A., Yeakum M., Bell M., Dedhia N.,		
RA	Pannell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,		
RA	Kitchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,		
RA	Mattisensen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,		
RA	Voicikert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,		
RA	Erman K. D., Terry N., Hartley N., Bent E., Johnson S.,		
RA	Langham S.-A., McCulligh B., Robben J., Grymopre B., Zimmermann W.,		
RA	Ranasperger U., Wedler H., Balke K., Wedler E., Peters S.,		
RA	Van Staveren M., Dirkse W., Mooljaan P., Klein Lankhorst R.,		
RA	Feldeneeger T., Bothe G., Rose M., Hauf J., Benneiser S., Hempel S.,		
RA	Weldpausch M., Lamberth S., Villarroel R., Gjeiljen J., Ardiles W.,		
RA	Berts O., Lemcke K., Kolev G., Mayer K.F.X., Rudd S., Schoot H.,		
RA	Schneeller C., Zaccaria P., Mewes H.-W., Bevan M., Franz P.F.,		
RT	"Sequence and analysis of chromosome 5 of the plant Arabidopsis		

```
RT thaliana.;
RL Nature 408:823-826 (2000).
CC -!- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
CC PATHWAY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -!- MISCELLANEOUS: CPR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
CC RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
CC COTYLEDON GROWTH IS IMPAIRED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MIL/RAF SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L08789; AAA32779.1; -.
CC EMBL; L08790; AAA32780.1; -.
CC EMBL; AL162506; CAB82938.1; -.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Serine/threonine-protein kinase; ATP-binding.
CC DOMAIN 65 69 POLY-GLY.
CC FT DOMAIN 135 141 POLY-GLY.
CC FT DOMAIN 551 809 PROTEIN_KINASE.
CC FT NP_BIND 557 565 ATP (BY SIMILARITY).
CC FT BINDING 578 578 ATP (BY SIMILARITY).
CC FT ACT_SITE 676 676 BY SIMILARITY.
CC FT MUTAGEN 596 596 E->K; IN CTRL-4; EXHIBITS ETHYLENE-
CC TREATED PHENOTYPE.
CC FT MUTAGEN 694 694 D->E; IN CTRL-1; EXHIBITS ETHYLENE-
CC TREATED PHENOTYPE.
CC FT SEQUENCE 821 AA; 90306 MW; 29223D3DCDCC15BC CRC64;
CC -----
Query Match 24.6%; Score 398; DB 1; Length 821;
Best Local Similarity 33.5%; Pred. No. 3.9e-27;
Matches 90; Conservative 60; Mismatches 107; Indels 12; Gaps 8;
QY 30 EIDYKIEVEEVVGRGAFGVVCKAKWRAKDVAIK---QIESESER-KAFIVELQQLSRVN 85
DB 545 DIPWCDLNTKEKIGAGSFGFVTHRAHGHGSDVAVKILMEQDFHAERVNEFLREVAIMKRLR 604
QY 86 HPNIVKLYGACLNPP--VCLVMEYAEAGGSLYNVLHGAEPYPYVTAAHAMSWCLQCQGVAY 143
DB 605 HPNIVLFMGAVTQPNLSIVTEVLSGSLYRLHLKSGAREQLDERRLSMAVDVAKGMY 664
QY 144 LHSMPQKALIHRLDKLPNLLVAGGTGVLKICDPG-TACDIOHTMTNNK--GSAANWAPEV 200
DB 665 LHNRP-PIVHRLDKLPNLLVDDKKYTV-KVCDPGLSLKASTFLSKSAAGTPENWAPEV 722
QY 201 FEGSNSEKCDVPSWGIILWEVITRKPKDEIGGPAFRIMWAVHNGTRPLPKNPKPIE 260
DB 723 LRDEPNSEKSDVYSFGVILLWEALATQPGWNL--NPAQVVAAGFKCKRLEIPRNLPQVA 781
QY 261 SLMTRGWSKDPSPORSPMEETVKIMTHLMR 289
DB 782 AITEGCTWNPWKRPSEATIMDLRLPLIK 810
RESULT 11
KYK1_DICDI
ID KYK1_DICDI STANDARD; PRT; 1584 AA.
```

```
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
DE protein kinase 1).
GN PYKA OR SPLA OR DPK1.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
CX NCBI_taxid=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JH10;
RX MEDLINE=97053827; PubMed=8898241;
RA Nuckolls G.H., Osherov N., Loomis W.F., Spudich J.A.;
RT "The Dictyostelium dual-specificity kinase spla is essential for
RT spore differentiation.";
RL Development 122:3295-3305 (1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
RT Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583 (1990).
CC -!- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -!- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32174; AAB41125.1; -.
CC EMBL; M33785; AAA33202.1; -.
CC PIR; A35670; A35670.
CC DictyDb; DD03010; pyka.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR001660; SAM.
CC InterPro; IPR003878; SPRY domain.
CC InterPro; IPR003877; SPRY receptor.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR001245; Tyr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00536; SAM; 1.
CC Pfam; PF00622; SPRY; 3.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00454; SAM; 1.
CC SMART; SM00449; SPRY; 3.
CC SMART; SM00221; STYK; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
CC PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE; PS0105; SAM_DOMAIN; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 596 600 POLY-ASN.
FT DOMAIN 808 811 POLY-PHE.
FT DOMAIN 1026 1029 POLY-SER.
```



```
FT VARSPLIC 224 245 MISSING (IN ISOFORM 2).
FT VARSPLIC 604 630 RPEGRPSLELLTIDELVECFEGR -> ESCLCRVAQD
FT LSXKNLIGSRF (IN ISOFORM 3).
FT CONFLICT 535 535 V -> E (IN REF. 3).
FT CONFLICT 550 553 FGVL -> YGIP (IN REF. 3).
FT CONFLICT 611 611 L -> F (IN REF. 2).
SQ SEQUENCE 630 AA; 73426 MW; 262640EB90D4A6D2 CRC64;

Query Match 22.6%; Score 365.5; DB 1; Length 630;
Best Local Similarity 34.3%; Pred. No. 2e-24;
Matches 92; Conservative 49; Mismatches 110; Indels 17; Gaps 9;

QY 30 EIDYKEIEVEVVGKGFAGVGVVCKAKWRAK-DVAIKQI-ESESERKAFIVELRQLSRVNH 87
DB 363 EINFSEITFMRELGSGLFGVVLGKWAQYKVAIKALREGAMCEEDFIEAKVMKLT 422
QY 88 NIVKLYGACL--NPVCLVMEYAEGLSYLVHGAEPILPYTYAAHMSWCLQCQGVAYLH 145
DB 423 KLVQLYGVCTQOKPIYIVTFMERGCLNLFRLQRQ--GHFSRDMLLSMCDVCEGMEYL- 479
QY 146 SMQPKALIHRLDKPPNLLVAGGTVLKICDFGTA----CDIQTHMTNNKGSAAWMAPEVF 201
DB 480 --ERNSTFIHRDLAARNCLVNEAG-VVKVSDFGMARVYLDQYTSSGAKFPVKWCPPEVF 536
QY 202 EGSNYSEKCDVFSWGIIILWEVITR-RKPFDEIGGPAFRIMVAVHNGTRPLIKNLPKPIE 260
DB 537 NYSRFSKSDVWSFGVLMEIFTEGRMPFEK--NTNVEVVTWTRGHRLHRLPKLATKYLY 594
QY 261 SLMTRCWSKDPSPQRSMEIIVKINTHLM 288
DB 595 EVMLRCWQERPEGRPSLEDLLRTIDELV 622

RESULT 13
TEC_HUMAN STANDARD; PRT; 631 AA.
AC P42680;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
GN TEC OR PSCTK4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=95019807; PubMed=7934162;
RA Sato K., Mano H., Ariyama T., Inazawa J., Yazaki Y., Hirai H.;
RT "Molecular cloning and analysis of the human Tec protein-tyrosine
RT kinase.";
RL Leukemia 8:1663-1672(1994).
CC -! CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -! SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -! TISSUE SPECIFICITY: HEMATOPOIETIC CELL LINES INCLUDING MYELOID,
CC B-, AND T-CELL LINEAGES.
CC -! SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
CC SUBFAMILY.
CC -! SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -! SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -! SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -! CAUTION: IT IS UNCERTAIN WHETHER MET-1 IS THE INITIATOR.
CC -! DATABASE: NAME=Atlas Genet. Cytogenet. Oncol. Haematol.;
CC WWW="http://www.infobiogen.fr/services/chronocare/Genes/TECID75.html".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D29767; BAA06171.1; -.
DR HSSP; Q06187; 1B55.
DR Genew; HGNC:11719; TEC.
DR MIM; 600583; -.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00402; TECBKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00003; PH DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation.
FT DOMAIN 4 111 PH.
FT DOMAIN 179 239 SH3.
FT DOMAIN 247 345 SH2.
FT DOMAIN 370 623 PROTEIN KINASE.
FT NP_BIND 376 384 ATP (BY SIMILARITY).
FT BINDING 398 398 ATP (BY SIMILARITY).
FT ACT_SITE 489 489 BY SIMILARITY.
FT MOD_RES 519 519 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 631 AA; 73629 MW; A55DECAF991A9022 CRC64;

Query Match 22.6%; Score 365.5; DB 1; Length 631;
Best Local Similarity 34.7%; Pred. No. 2e-24;
Matches 94; Conservative 47; Mismatches 107; Indels 23; Gaps 11;

QY 30 EIDYKEIEVEVVGKGFAGVGVVCKAKWRAK-DVAIKQI-ESESERKAFIVELRQLSRVNH 87
DB 364 EINFSEITFMRELGSGLFGVVLGKWAQYKVAIKALREGAMCEEDFIEAKVMKLT 423
QY 88 NIVKLYGACL--NPVCLVMEYAEGLSYLVHGAEPILPYTYAAHMSWCLQCQGVAYLH 145
DB 424 KLVQLYGVCTQOKPIYIVTFMERGCLNLFRLQRQ--GHFSRDVLLSMCDVCEGMEYL- 480
QY 146 SMQPKALIHRLDKPPNLLVAGGTVLKICDFGTA----CDIQTHMTNNKGSAAWMAPEVF 201
DB 481 --ERNSTFIHRDLAARNCLVSEAG-VVKVSDFGMARVLDQYTSSGAKFPVKWCPPEVF 537
QY 202 EGSNYSEKCDVFSWGIIILWEVITR-RKPFDEIGGPAFRIMVAVHNGTR---PPLIKNLPK 257
DB 538 NYSRFSKSDVWSFGVLMEVTEGRMPFEKYN--YEVVTWTRGHRLYQPKLASNY-- 593
QY 258 PIESLMTRCWSKDPSPQRSMEIIVKINTHLM 288
DB 594 -YEVMLRCWQERPEGRPSLEDLLRTIDELV 623
```

```

RESULT 14
ITK_HUMAN STANDARD; PRT; 620 AA.
ID ITK_HUMAN
AC 008881;
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific
kinase) (Tyrosine-protein kinase Ltk) (Kinase EMT).
GN ITK OR LTK OR EMT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93279365; PubMed=8504851;
RA Tanaka N., Asao H., Ohnani K., Nakamura M., Sugamura K.;
RT "A novel human tyrosine kinase gene inducible in T cells by
interleukin 2.";
RL FEBS Lett. 324:1-5(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymus; PubMed=8364206;
RA Gibson S., Leung B., Squire J.A., Hill M., Arima N., Goss P.,
RA Hogg D., Mills G.B.;
RT "Identification, cloning, and characterization of a novel human
T-cell-specific tyrosine kinase located at the hematopoietin complex
on chromosome 5q.";
RL Blood 82:1561-1572(1993).
CC -1- FUNCTION: PLAYS A ROLE IN T CELL PROLIFERATION AND
DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
tyrosine phosphate.
CC -1- SUBCELLULAR LOCATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CD28, CD2) IN T-CELLS (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: T-CELL LINES AND NATURAL KILLER CELL LINES.
CC -1- INDUCTION: BY INTERLEUKIN-2.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; D13720; BAA02873.1; -
DR EMBL; L10717; AAA36748.1; -
DR EMBL; S65186; AAB28072.2; -
DR PIR; S33253; S33253.
DR HSSP; Q03526; 1AMU.
DR Genew; HGNC:6171; ITK.
DR MIM; 186973; -.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00401; SH2DOMAIN.

```

```

DR PRINTS; PR00402; TECBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_Pkinase; 1.
DR Prodom; PD000066; SH3; 1.
DR Prodom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; Tyrc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
DR Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain.
KW DOMAIN
FT DOMAIN 171 231 SH3.
FT DOMAIN 239 338 SH2.
FT DOMAIN 363 615 PROTEIN_KINASE.
FT NP_BIND 369 377 ATP (BY SIMILARITY).
FT BINDING 391 391 ATP (BY SIMILARITY).
FT ACT_SITE 482 482 BY SIMILARITY.
FT MOD_RES 512 512 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 331 331 V -> W (IN REF. 2; AAB28072).
SQ SEQUENCE 620 AA; 71831 MW; DAE396BD2309319D CMC64;

Query Match 22.2%; Score 358.5; DB 1; Length 620;
Best local Similarity 33.5%; Pred. No. 8.2e-24;
Matches 91; Conservative 43; Mismatches 109; Indels 29; Gaps 11;

QY 31 IYKKEIEVEVIGRGAFGVYCKAKWRAD-VAIKOI-ESESRRKAFIVELRLQSLVHNP 88
DB 358 IDPELTVEQETGSGQFGVHLGWLNDKVAIKITIRGAMSEEDFIEAEVYMKLSHK 417
QY 89 IYLYAGACLN--PVCIVMEYAEAGSLYVNLHGAEPDLYPTAAAHAMSWCICQSGVAYLHS 146
DB 418 IYGLYVCELEQAPICLVFEFMHGCISDLRTQRGL--FAETLLGMCILDVEGNAVYL-- 473
QY 147 MQPKALIRHDLKPNLLVAGTGLKICDFG---TACDIQTHMTNNGSAAMAPVEYF 202
DB 474 -EBAVCVHRDLAARN-CLVGENOVIKVSDFGMTFRVLDDQYTSSTGTFKPVYKASPEVFS 531
QY 203 GSNYSKCVFPMGILWAVITRRK-PPDEIGCPAFRIMAVHNGTRPPLIKNLPK--- 258
DB 532 FSRYSKSDVWSFGVLMWVFESEGIPEYENRNS--EVEDISTGFR-----LYKPRLA 583
QY 259 ---IESLMTRCWSKDPQSPSMEETIKIMTHL 287
DB 584 STHVYQIMNHCKKERPEDRPAFSRLRLRLAET 615

RESULT 15
ITK_MOUSE STANDARD; PRT; 625 AA.
ID ITK_MOUSE
AC 003526;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase ITK/TSK (EC 2.7.1.112) (T-cell-specific
kinase) (Il-2-inducible T-cell kinase) (Kinase EMT) (Kinase TLK).
GN ITK OR TSK OR EMT OR TLK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymocytes; PubMed=1280821;
RX MEDLINE=93087493; PubMed=1280821;
RA Siliciano J.D., Morrow T.A., Desiderio S.V.;

```

RT "itk, a T-cell-specific tyrosine kinase gene inducible by interleukin
RL 2.";
RN Proc. Natl. Acad. Sci. U.S.A. 89:11194-11198(1992).
RN (2)
RC SEQUENCE FROM N.A.
RP TISSUE=Thymocytes;
RX MEDLINE=93133848; PubMed=8421704;
RA Heycock S.D., Berg L.J.;
RT "Developmental regulation of a murine T-cell-specific tyrosine kinase
RL gene, Tsk.";
RN Proc. Natl. Acad. Sci. U.S.A. 90:669-673(1993).
RN (3)
RC SEQUENCE FROM N.A.
RP STRAIN=CBA/J; TISSUE=Mast cells;
RX MEDLINE=93236578; PubMed=8476425;
RA Yamada N., Kawakami Y., Kimura H., Fukumachi H., Baier G.,
RL Altman A., Kato T., Inagaki Y., Kawakami T.;
RT "Structure and expression of novel protein-tyrosine kinases, Emb and
RL Emt, in hematopoietic cells.";
RN Biochem. Biophys. Res. Commun. 192:231-240(1993).
RN (4)
RC SEQUENCE FROM N.A.
RP Ogata M., Sawada M., Fujiwara H., Hamaoka T.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RN (5)
RC STRUCTURE BY NMR OF 160-236.
RX MEDLINE=97138229; PubMed=8985255;
RA Andreotti A.H., Bunnell S.C., Feng S., Berg L.J., Schreiber S.L.;
RT "Regulatory intramolecular association in a tyrosine kinase of the
RL Tec family.";
RN Nature 385:93-97(1997).
RN (6)
RC CHARACTERIZATION.
RX MEDLINE=95023908; PubMed=7524075;
RA August A., Gibson S., Kawakami Y., Kawakami T., Mills G.B., Dupont B.;
RT "CD28 is associated with and induces the immediate tyrosine
RL phosphorylation and activation of the Tec family kinase ITK/EMT in
RL the human Jurkat leukemic T-cell line.";
RN Proc. Natl. Acad. Sci. U.S.A. 91:9347-9351(1994).
RN (7)
RC CHARACTERIZATION.
RX MEDLINE=97098950; PubMed=8943565;
RA King P.D., Sadra A., Han A., Liu X.-R., Sunder-Plassmann R.,
RA Reinherz E.L., Dupont B.;
RT "CD2 signaling in T cells involves tyrosine phosphorylation and
RL activation of the Tec family kinase, EMT/ITK/TSK.";
RN Int. Immunol. 8:1707-1714(1996).
RN (8)
RC CHARACTERIZATION.
RX MEDLINE=20040393; PubMed=10570288;
RA Ching K.A., Kawakami Y., Kawakami T., Tsoukas C.D.;
RT "Emt/Itk associates with activated TCR complexes: role of the
RL pleckstrin homology domain.";
RN J. Immunol. 163:6006-6013(1999).
RN (9)
CC -!- FUNCTION: MAY PLAY A ROLE IN T-CELL DEVELOPMENT, POTENTIALLY IN
CC THYMIC SELECTION.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- SUBUNIT: LIGATION OF CD2, TCR OR CD28 INDUCES ACTIVATION AND
CC TYROSINE PHOSPHORYLATION OF ITK.
CC -!- SUBCELLULAR LOCALIZATION: LOCALIZES TO CELL SURFACE RECEPTORS IN THE
CC PLASMA MEMBRANE AFTER STIMULATION WITH RESPECTIVE RECEPTORS (TCR,
CC CD28, CD2) IN T-CELLS.
CC -!- TISSUE SPECIFICITY: IS DETECTED IN THE THYMUS, LYMPH NODE AND VERY
CC FAINTLY IN THE SPLEEN, BUT IS NOT DETECTED IN THE LIVER, LUNG,
CC KIDNEY, HEART, BRAIN, INTESTINE OR TESTIS. EXPRESSED IN T-
CC LYMPHOCYTES AND MAST CELLS. IT MAY ALSO BE EXPRESSED IN NATURAL
CC KILLER CELLS.
CC -!- DEVELOPMENTAL STAGE: IS PRESENT IN THE FETAL THYMUS AS EARLY AS
CC DAY 14 OF GESTATION. THE LEVELS ARE 5- TO 10-FOLD HIGHER IN
CC THYMOCYTES THAN IN PERIPHERAL T CELLS, AND INCREASE IN THE THYMUS
CC DURING DEVELOPMENT FROM NEONATE TO ADULT.
CC -!- INDUCTION: BY INTERLEUKIN-2.

CC -!- DOMAIN: THE PH DOMAIN MEDIATES TARGETING OF ITK/TSK AND IS
CC INDISPENSIBLE FOR THE ACTIVATION THROUGH TCR/CD3.
CC -!- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
CC SUBFAMILY.
CC -!- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L00619; AAA39337.1; -;
DR EMBL; L05631; AAA40518.1; -;
DR EMBL; L10628; -; NOT_ANNOTATED_CDS.
DR EMBL; D14042; BAA03129.1; -;
DR PIR; A43030; A43030.
DR PDB; 1AWJ; 14-JAN-98.
DR MGD; MGI:96621; Itk.
DR InterPro; IPR001562; BTK.
DR InterPro; IPR007719; Euk_pkinase.
DR InterPro; IPR001849; PH.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00017; SH2; 1.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00402; TECPTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE_TVR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain; 3D-structure.
FT DOMAIN 4 117
FT DOMAIN 177 237
FT DOMAIN 245 343
FT DOMAIN 368 620
FT NP_BIND 374 382
FT BINDING 396 396
FT ACT_SITE 487 487
FT MOD_RES 517 517
FT CONFLICT 82 87
FT CONFLICT 535 535
FT CONFLICT 540 540
SQ SEQUENCE 625 AA; 72291 MW; F7A4A18A8A1AADDC CRC64;
Query Match 22.0%; Score 356; DB 1; Length 625;
Best Local Similarity 32.4%; Pred. No. 1.4e-210;
Matches 93; Conservative 44; Mismatches 110; Indels 40; Gaps 12;
QY 16 GMEIAPSQVLNFEEDIDYKEIEVEVVGRGAFGVCKAKWRAD-VAIKOI-ESESERKA 73
Db 359 GKWVIOQSELTFVQEI-----GSGQGLVHLGVLNKKDKVAIKTIQEGAMSEED 407

```

QY 74 FIVEIROLSRHNHNIYKLYACXCN--PVCIMWEYAGSGIYUNYHGAEDLPYXTAHAM 111
D 408 FIEEVEVMKSHKIKQYLYVCEIQAPICLVEEFMHGCSIDILRSQRI--FAAETTL 4655
QY 132 SWCIQSGQVAYIYHSMOPKALIHEDLXPMLLLVAGGTVLKICDFG---TACDIQHHMT 1877
D 466 GMCLDVEGMAAYL--EKACVIRHDLAARN-CLVGENQVLIKSPFGMTRFVLDDQYLSST 521
QY 188 NNRKSAAAMAPVEVEGSGNSYKCVFSGMGIIMVYITRRK-PFEDIGCPAFRIMAWANG 2466
D 522 GTRKPPVKAASEVEFSFRYSKSDVMSFGVLMVMEVFEBGKIPEYENRNS--EVEDISTG 579
QY 247 TRPELIKMLPP-----IESIMTRCSKDPQSPQEMEEIYKIMTHL 287
D 580 FR-----LYKPELASCHVQYIMNHCKEKEDRPPPSQILSQAETI 620

```

Search completed: December 10, 2002, 03:48:49
Job time : 45 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:43:10 ; Search time 96 Seconds
(Without alignments)
650.336 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615
Sequence: 1 MSTAASASSSSSSAGEMIE.....MTIMKRYPPGADEPLQPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

SPRREMBL_21: *
1: sp_archaea: *
2: sp_bacteria: *
3: sp_fungi: *
4: sp_human: *
5: sp_invertebrate: *
6: sp_mammal: *
7: sp_mmc: *
8: sp_organelle: *
9: sp_phase: *
10: sp_plant: *
11: sp_rodent: *
12: sp_virus: *
13: sp_vertebrate: *
14: sp_unclassified: *
15: sp_virus: *
16: sp_bacteriap: *
17: sp_archaeap: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1615	100.0	491	Q9NZ70	Q9NZ70 homo sapien
2	1609	99.6	606	Q9Z3A8	Q9Z3A8 mus musculu
3	1490.5	92.3	616	13 O73613	O73613 xenopus lae
4	1428	88.4	478	4 Q9NTR4	Q9NTR4 homo sapien
5	1428	88.4	539	4 Q9NTR1	Q9NTR1 homo sapien
6	1428	88.4	566	4 Q9NTR2	Q9NTR2 homo sapien
7	813	50.3	678	5 Q9V3O6	Q9V3O6 drosophila
8	505	31.3	252	5 Q9VCV0	Q9VCV0 drosophila
9	442	27.4	1066	4 Q9H2N5	Q9H2N5 homo sapien
10	434.5	26.9	411	10 Q9ZQ31	Q9ZQ31 arabidopsis
11	434	26.9	412	10 Q9M085	Q9M085 arabidopsis
12	434	26.9	454	11 Q9ESL3	Q9ESL3 mus musculu
13	434	26.9	802	11 Q9ESL4	Q9ESL4 mus musculu
14	433	26.8	371	13 Q9QZY8	Q9QZY8 brachydanio
15	433	26.8	455	4 Q9HCC4	Q9HCC4 homo sapien
16	433	26.8	800	4 Q9HDD2	Q9HDD2 homo sapien

17	433	26.8	800	4 Q9HCC5	Q9HCC5 homo sapien
18	433	26.8	800	4 Q9NYL2	Q9NYL2 homo sapien
19	433	26.8	800	4 Q9NYE9	Q9NYE9 homo sapien
20	432	26.7	564	4 Q9H1Y7	Q9H1Y7 homo sapien
21	430	26.6	422	5 Q23846	Q23846 dictyosteli
22	430	26.6	1001	11 Q8VDC6	Q8VDC6 mus musculu
23	430	26.6	1338	5 Q23927	Q23927 dictyosteli
24	428	26.5	570	4 Q8MWN2	Q8MWN2 homo sapien
25	428	26.5	1036	4 Q8MWN1	Q8MWN1 homo sapien
26	427.5	26.5	391	10 Q22100	Q22100 arabidopsis
27	427.5	26.5	1020	5 Q9W3I3	Q9W3I3 drosophila
28	427.5	26.5	1148	5 Q9SVF6	Q9SVF6 drosophila
29	427.5	26.5	1161	5 Q9SUN8	Q9SUN8 drosophila
30	421.5	26.1	357	10 Q9S1M8	Q9S1M8 arabidopsis
31	421	26.1	850	11 Q9J1T5	Q9J1T5 mus musculu
32	419	25.9	847	4 Q16584	Q16584 homo sapien
33	418.5	25.9	416	10 Q94C42	Q94C42 triticum ae
34	416	25.8	740	5 Q21982	Q21982 caenorhabdi
35	412	25.5	483	10 Q8RY96	Q8RY96 arabidopsis
36	412	25.5	570	10 Q8RWL6	Q8RWL6 arabidopsis
37	411.5	25.5	546	10 Q2S558	Q2S558 arabidopsis
38	411.5	25.5	553	10 Q81808	Q81808 arabidopsis
39	411	25.4	1030	10 Q9C9U5	Q9C9U5 arabidopsis
40	410	25.4	835	4 Q9Y2V6	Q9Y2V6 homo sapien
41	408	25.3	411	10 Q9AWA6	Q9AWA6 atachis hyp
42	406.5	25.2	370	10 Q9S7D5	Q9S7D5 arabidopsis
43	404	25.0	847	10 Q93XU9	Q93XU9 rosa hybrid
44	401.5	24.9	527	5 Q9B1Z5	Q9B1Z5 dictyosteli
45	401	24.8	859	4 Q8WY25	Q8WY25 homo sapien

ALIGNMENTS

RESULT 1
Q9NZ70 PRELIMINARY; PRT; 491 AA.
ID Q9NZ70 Q9NTR3;
AC Q9NZ70; Q9NTR3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE TGF beta-activated kinase splice variant d (D0154G14.1.4)
DE (mitogen-activated protein kinase kinase 7 (TGF-beta activated kinase 1d (TAK1)))
DE kinase 1d (TAK1))
GN TAK1 OR MAP3K7
GN Home sapiens (Human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568765; PubMed=1118615;
RA Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;
RT "Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1.";
RL Biochim. Biophys. Acta 1517:46-52 (2000).
RN [2]
RP SEQUENCE OF 41-491 FROM N.A.
RA Tracey A.;
RC Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF218074; AAF27652.1; -;
DR EMBL; AL121964; CAB87605.1; -;
DR HSSP; P06631; IADS.
DR HSSP; P12931; IFMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; Sry_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase_1.
DR PRINTS; PR00109; TRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.

```
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 491 AA; 53739 MW; B7D8832E286A99C5 CRC64;

Query Match 100.0%; Score 1615; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 3.5e-148; Indels 0; Gaps 0;
Matches 303; Conservative 0; Mismatches 0;

QY 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
DB 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60

QY 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
DB 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120

QY 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADEPLOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 2
Q923A8 PRELIMINARY; PRT; 606 AA.
ID Q923A8;
AC Q923A8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5989).
GN MAP3K7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006665; AAH06665.1; -.
DR MGD; MGI:1346877; Map3k7.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Transferase.
SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Query Match 99.6%; Score 1609; DB 11; Length 606;
Best Local Similarity 99.7%; Pred. No. 1.8e-147;
Matches 302; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60
DB 1 MSTASASSSSSSSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDV 60

QY 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
```

```
DB 61 AIKQIESESERKAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAE 120
QY 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180
DB 121 PLPYITAAHAMSWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTAC 180

QY 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240
DB 181 DIQTHMTNKGSAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIM 240

QY 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADEPLOY 300
DB 241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMRYFFGADEPLOY 300

QY 301 PCQ 303
DB 301 PCQ 303

RESULT 3
Q73613 PRELIMINARY; PRT; 616 AA.
ID Q73613;
AC Q73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAKI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
SEQUENCE FROM N.A.
RP MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAKI and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBL J. 17:1019-1028 (1998).
CC -/- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U92030; AAC14008.1; -.
DR HSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Query Match 92.3%; Score 1490.5; DB 13; Length 616;
Best Local Similarity 94.9%; Pred. No. 5.7e-136;
Matches 277; Conservative 6; Mismatches 8; Indels 1; Gaps 1;

QY 12 SSSAGEMIEAPSOVLNFEIDYKEIEVEEVVGRGAFGVCKAKWRAKDVAIKQIESESER 71
DB 2 SATSAEMIETP-PVLNFEIDYKEIEVEEVVGRGTFGVCKAKWRAKDVAIKQIESESER 60

QY 72 KAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPLPYITAAHAM 131
DB 61 KAFIVELRQLSRVNHPIVVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPLPYITAAHAM 120

QY 132 SWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKG 191
DB 121 SWCLOCSOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNKG 180

QY 192 SAAMWAPVFEFGSNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMAVHNGTRPPL 251
```

```

Db 181 SAAMMAPEVEGSGNSYSEKCVFSWGIIIMVITRRKPFDEIGGAPAFRIMAVANGTRPPL 240
QY 252 IKULPRIESIMTRCWSKDSQSPSMEEIYKIMTHLMRYFPFGADEPLQYPCQ 303
Db 241 IKULPRIESIMTRCWSKDSQSPSMEEIYKIMTHLMQYFPFGADVSLQYPCQ 292

RESULT 4
Q9NTR4 PRELIMINARY; PRT; 478 AA.
ID Q9NTR4
AC Q9NTR4;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.3 (Mitogen-activated protein kinase kinase 7
  (TGF-beta activated kinase 1c (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87604.1; -.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Kinase; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 478 AA; 52482 MW; 177CC8CFA8DBF8 CRC64;

Query Match 88.4%; Score 1428; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 4, 8e-130; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 41 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 100
Db 1 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 60

QY 101 CLVMEYABGSGLYNVLHGAEPLPYTTAAHAMSWCLOCSQGVAYIHSNPKALIHRLDKP 160
Db 61 CLVMEYABGSGLYNVLHGAEPLPYTTAAHAMSWCLOCSQGVAYIHSNPKALIHRLDKP 120

QY 161 NLLVAGGTVLKICDPGACDIQTHMTNNKGSAAWMAPEVEGSGNSYSEKCVFSWGIIIM 220
Db 121 NLLVAGGTVLKICDPGACDIQTHMTNNKGSAAWMAPEVEGSGNSYSEKCVFSWGIIIM 180

QY 221 EVITRRKPFDEIGGAPAFRIMAVANGTRPPLIKULPKPIESIMTRCWSKDSQSPSMEEI 280
Db 181 EVITRRKPFDEIGGAPAFRIMAVANGTRPPLIKULPKPIESIMTRCWSKDSQSPSMEEI 240

QY 281 VKIMTHLMRYFPFGADEPLQYPCQ 303
Db 241 VKIMTHLMRYFPFGADEPLQYPCQ 263

RESULT 5
Q9NTR1 PRELIMINARY; PRT; 539 AA.
ID Q9NTR1
AC Q9NTR1;

```

```

DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.1 (Mitogen-activated protein kinase kinase 7
  (TGF-beta activated kinase 1a (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87607.1; -.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Kinase; Serine/threonine-protein kinase.
FT NON_TER 1
SQ SEQUENCE 539 AA; 60006 MW; E6183F553CCTF324 CRC64;

Query Match 88.4%; Score 1428; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 5, 6e-130; Indels 0; Gaps 0;
Matches 263; Conservative 0; Mismatches 0;

QY 41 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 100
Db 1 VVGRGAFGVCCAKWRADVAIKQISESESKAFIVELRQLSRVNHPIVLYGACINPV 60

QY 101 CLVMEYABGSGLYNVLHGAEPLPYTTAAHAMSWCLOCSQGVAYIHSNPKALIHRLDKP 160
Db 61 CLVMEYABGSGLYNVLHGAEPLPYTTAAHAMSWCLOCSQGVAYIHSNPKALIHRLDKP 120

QY 161 NLLVAGGTVLKICDPGACDIQTHMTNNKGSAAWMAPEVEGSGNSYSEKCVFSWGIIIM 220
Db 121 NLLVAGGTVLKICDPGACDIQTHMTNNKGSAAWMAPEVEGSGNSYSEKCVFSWGIIIM 180

QY 221 EVITRRKPFDEIGGAPAFRIMAVANGTRPPLIKULPKPIESIMTRCWSKDSQSPSMEEI 280
Db 181 EVITRRKPFDEIGGAPAFRIMAVANGTRPPLIKULPKPIESIMTRCWSKDSQSPSMEEI 240

QY 281 VKIMTHLMRYFPFGADEPLQYPCQ 303
Db 241 VKIMTHLMRYFPFGADEPLQYPCQ 263

RESULT 6
Q9NTR2 PRELIMINARY; PRT; 566 AA.
ID Q9NTR2
AC Q9NTR2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE D154G14.1.2 (Mitogen-activated protein kinase kinase 7
  (TGF-beta activated kinase 1b (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

Tracey A.;
 RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL121964; CAB87606.1; -;
 DR HSSP; P12931; 1FMK.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Kinase; Serine/threonine-protein kinase.
 FT NON_TER 1
 ST SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;
 Query Match 88.4%; Score 1428; DB 4; Length 566;
 Best Local Similarity 100.0%; Pred. No. 6e-130;
 Matches 263; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 41 VVGRGAFGVVCKAKWRAKDVAIKQISESERKAFIVELRQLSRVNHFNIVKLYGACLNVPV 100
 DB 1 VVGRGAFGVVCKAKWRAKDVAIKQISESERKAFIVELRQLSRVNHFNIVKLYGACLNVPV 60
 QY 101 CLVMEYAGGSLNVLHGAEPLPYVTAAHAMSWCLOCSQGVAYLHSMQKALIHRLDKPP 160
 DB 61 CLVMEYAGGSLNVLHGAEPLPYVTAAHAMSWCLOCSQGVAYLHSMQKALIHRLDKPP 120
 QY 161 NLLLVAGGTVLKICDFGTACDIOTHTNNKGSAAWMAPEVFGNSYSEKCDVFSWGIILW 220
 DB 121 NLLLVAGGTVLKICDFGTACDIOTHTNNKGSAAWMAPEVFGNSYSEKCDVFSWGIILW 180
 QY 221 EVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLPKPIESLMTWCSDKPSRPSMEI 280
 DB 181 EVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNLPKPIESLMTWCSDKPSRPSMEI 240
 QY 281 VKIMTHLMRYFPGADEPLQVPCQ 303
 DB 241 VKIMTHLMRYFPGADEPLQVPCQ 263
 RESULT 7
 Q9V3Q6 PRELIMINARY; PRT; 678 AA.
 AC Q9V3Q6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative TAK1 protein (CG1388 protein) (LD42274P).
 GN TAK1 OR CG1388 OR CG18492.
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxId=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Framkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cayley S., Dahlke C., Davenport L.B., Davies P., M.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 [2]
 SEQUENCE FROM N.A.
 RP Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
 RA O'Connor M.B., Shibuya H., Ueno N.;
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuncio J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AE003571; AAP50895.1; -;
 DR EMBL; AF199466; AAF06815.1; -;
 DR EMBL; AY051953; AAK93377.1; -;
 DR HSSP; P08631; IAD5.
 DR FlyBase; FBgn0026323; Tak1.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR001990; Granin.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKC; 1.
 DR PROSITE; PS00422; GRANINS_1; UNKNOWN_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Hypothetical protein; Kinase;
 KW Serine/threonine-protein kinase; Transferase.
 ST SEQUENCE 678 AA; 75674 MW; 87EBA80CDB8CDE45 CRC64;
 Query Match 50.3%; Score 813; DB 5; Length 678;
 Best Local Similarity 54.0%; Pred. No. 3.8e-70;
 Matches 154; Conservative 45; Mismatches 82; Indels 4; Gaps 3;
 QY 18 MIEAPSQLNPEEDIDYKEIEVEEVVGAGVGVCYKAKWRAKDVAIKQISESERKAFIVE 77
 DB 1 MATASLDALQAAVYDFSEITLREKVGHSYGVVCKAVWRDKLVAVKEFFASAQKQIEKE 60

```

Qy 78 LRQSRVNHRIIVLYG--ACINPVCUMVEYABEGSLVYNLHGAEPILPYTAHAMS WC L 135
Db 61 VKQISRVKHHPIIALHGISSYQOATYILIMEFABEGSLHNLHAG-KVKPAYSLAHAMS MAR 119
Qy 136 QCSGVAVYLSMOPKALIHRLDKPNNLLVAGGTVLKICDFTACDIQTHMTNNKSAAM 195
Db 120 QCAAGLAVYLAAMPKPIIHHDVKKPILNLLLNKGRNLKICDFTGVAADKSTMTNNKSGAAM 179
Qy 196 MAPEVFEKSYSEKCDVFSWGIILMEVITRRKPEDEIGGPAFRIMAVNHGTRPPLIKNL 255
Db 180 MAPEVFEKSYTEKCDIFSNVAVLMEVLSRKQPKGIDN-AYTLQMKVIYKGEHPPLITTC 238
Qy 256 PKPIESLMTKWSDBSPQSPMEIVIKIMTHLMKRYFPGADEPLQY 300
Db 239 PKRIEDLMTACWKTPBDRPSMOYIVGMHEIVKDYGAUKALEY 283

RESULT 8
Q9VCV0 PRELIMINARY; PRT; 252 AA.
ID Q9VCV0;
AC Q9VCV0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Phryganea; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gccayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champagne M., Pfeiffer B.D.,
RA Men K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Adair J.F., Agbayan A., An H.-U., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolehakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Maye A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glaser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegum C.,
RA Jajal M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshirel A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleby J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schaefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton W., Strong R., Sun E.,
RA Svaykaskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao O., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).

```

```

CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AE003740; AAF56055.1; -.
DR FlyBase; FBgn0039015; CG4803.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 252 AA; 28853 MW; 8E006F2EABBE8D17 CRC64;

Query Match 31.3%; Score 505; DB 5; Length 252;
Best Local Similarity 40.6%; Pred. No. 8.2e-41;
Matches 99; Conservative 48; Mismatches 71; Indels 26; Gaps 4;

Qy 29 EIDDKTEIEVERVGRGAFGVYCAKAKVAIAIKQISESEKAFIVELRQLSRVHNP 88
Db 6 EGVPEEIIQTKELIGTGYGSGYRAVMNRRIALKRIREGCEDKIEREIVQLTKASHVN 65
Qy 89 IYKLYGACLNPCV--LVMEYAEGLVYVHGAEPILPYTAHAMS WC LQCSGVAVYLS 146
Db 66 IVELYGTSRHEGCALLMEFVDSLSFLLH-AKSKPSYSHAHAFNMAHQIAQGIAYLHG 124
Qy 147 MQPKALIHRLDKPNNLLVAGGTVLKICDFTACDIQTHMTNNKSAAMAPE----- 199
Db 125 MQPKAVIHRDIPKNNLLVAGGTVLKICDFTACDIQTHMTNNKSAAMAPE----- 184
Qy 200 -----VFSGSYSEKCDVFSWGIILMEVITRRKPEDEIGGPAFRIMAVNH 244
Db 185 KSNRIIINOPTGFOVKVLOGNPKDEKCDVYSWAIWFWEILSRKEPEQY-NTLFELYMAIN 243
Qy 245 NGTR 248
Db 244 EGKR 247

RESULT 9
Q9H2N5 PRELIMINARY; PRT; 1066 AA.
ID Q9H2N5;
AC Q9H2N5;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Mixed lineage kinase MLK1 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA McNeae J.U., Dower S.K., Guesdon F.;
RT "cDNA sequence and gene organisation of mixed lineage kinase 1.";
RL Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL; AF251442; AAG44591.1; -.
DR HSRF; P29355; ISEM.
DR InterPro; IPR000719; Euk_Pkinase.
DR InterPro; IPR001452; Ser_thr_pkinase.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00018; SH3; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR ProDom; PD000066; SH3; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

```


SQ SEQUENCE 412 AA; 46083 MW; 644F35A90210D488 CRC64;
Query Match 26.9%; Score 434; DB 10; Length 412;
Best Local Similarity 36.8%; Pred. No. 1.2e-33;
Matches 103; Conservative 56; Mismatches 89; Indels 32; Gaps 11;
QY 25 VLNEE--IDYKEIEVEEVVGRGAFGVCCAKMRKAVAIKQIE-----SESERKAF 74
DB 118 LVNVEEHTIDIRKILHMPAFQAQAFGLYRGTYNGEDVAIKLERSDSNDEKAQALEQCF 177
QY 75 IVEIRQLSRVNPVIVLYGACLNVP--CLVMEYAEGGSLYNVL----HGAELPYPTTAA 128
DB 178 QQEVSMIAFLGHPNIVPTGACIKPVMWCIVTEYAKGSVRQFTTRKQNRAPVPLKT---- 233
QY 129 HANWSCLOCSQGVAYVLSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTH-M 186
DB 234 -AVWQALDVARGMAYVHE---RNFIRHDLKSDNLLISADSI-KIADFGYARLEVOQEGM 288
QY 187 TNNKGAAMAAPEVEEGSNYSEKCVFSWGIIMEVITRRKPEDEIGG--PAFRIMWAVH 244
DB 289 TPETGYRWMAPEWIIQHRPYTQKVDVYSPGIVLWELITGLLPQNMTAVOAAFAV---VN 345
QY 245 NGTRPPLIKULPKRIEISLMTRCWSKDSQRPSEIYKIM 284
DB 346 RGVPRPTVADCLPVLGEIMTRCWDADPEVRPCEAFIYNLL 385
RESULT 12
Q9ESL3 PRELIMINARY; PRT; 454 AA.
AC Q9ESL3; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE MLTK-beta.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I.; Adachi M.; Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049732; BAB16443.1; -.
DR HSSP; P12931; IFMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00219; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 454 AA; 51366 MW; 35C2F8C0D729D9395 CRC64;
Query Match 26.9%; Score 434; DB 11; Length 454;
Best Local Similarity 36.7%; Pred. No. 1.4e-33;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;
QY 27 NFEIIDIKEIEVEEVVGRGAFGVCCAKM--RAKDVAIK--QIESSEKRAFIIVEIRQL 81
DB 7 SFVQIKFDLDLQFFENCGSGFSGYRAKWTISQDEKAVAKLKLIKKEAE-----IL 57

QY 82 SRVNHPIVTKLYGACLNIP--VCLVMEYAEGGSLYNVLHG--AEPLPYTTAAHANSWCLOQC 137
DB 58 SVLSHRNIIIQFVGLLEPPNIGVIEYASLSGLYDINSNSSEEN---DMEHIMTWADTV 114
QY 138 SQGVAYVLSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTAQDIQTHMTNKK--GSAAM 195
DB 115 AKGMHYLHMEAFVVKVIRHDLKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFFP 172
QY 196 MAPEVEEGSNYSEKCVFSWGIIMEVITRRKPEDEIGGPAFRIMW-AVNHGTRPPLIKN 254
DB 173 MAPEVIQSLPVSSETDITYSYGVVLMELTRVPPFKGLBG--LQVAMLVVEKNERLLTIPS 230
QY 255 LKPIESLMTRCWSKDSQRPSEIYKIM 284
DB 231 CPRSFAPLHQCWEADAKRPSFKQIISIL 260
RESULT 13
Q9ESL4 PRELIMINARY; PRT; 802 AA.
AC Q9ESL4; 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MLTK alpha.
GN ZAK OR MLTK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21264927; Pubmed=11042189;
RA Gotoh I.; Adachi M.; Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL J. Biol. Chem. 276:4276-4286(2001).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049733; BAB16442.1; -.
DR HSSP; P12931; IFMK.
DR MGD; MGI:1931274; ZAK.
DR InterPro; IPR000194; Arpase a/bcentre.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PR00109; TYRKINASE.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYKc; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00152; ARPASE_ALPHA_BETA; UNKNOWN_1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 802 AA; 91719 MW; D43IDF8F312A43CC CRC64;
Query Match 26.9%; Score 434; DB 11; Length 802;
Best Local Similarity 36.7%; Pred. No. 3e-33;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;
QY 27 NFEIIDIKEIEVEEVVGRGAFGVCCAKM--RAKDVAIK--QIESSEKRAFIIVEIRQL 81
DB 7 SFVQIKFDLDLQFFENCGSGFSGYRAKWTISQDEKAVAKLKLIKKEAE-----IL 57
QY 82 SRVNHPIVTKLYGACLNIP--VCLVMEYAEGGSLYNVLHG--AEPLPYTTAAHANSWCLOQC 137
DB 58 SVLSHRNIIIQFVGLLEPPNIGVIEYASLSGLYDINSNSSEEM---DMEHIMTWADTV 114

```
QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172
QY 196 MAPEVPEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLITPSS 230
QY 255 LPKPIESLMTKRCWSDPSQSPSMEIEIVKIM 284
Db 231 CPRSFAELLHQWEADAKRPSFKQIISIL 260

RESULT 14
Q90ZY8 PRELIMINARY; PRT; 371 AA.
AC Q90ZY8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein kinase Npk.
OS Brachydanio rerio (Zebrafish) (Zebra daniel).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RA Chou C.-M., Lee I.-L., Leu J.-H., Huang C.-J.;
RT "A novel protein kinase, znpk, from the zebrafish.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF265343; AAKS2416.1; -.
DR InterPro; IPR000719; Ser_thr_kinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN_1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 371 AA; 42456 MW; 9B918B8A8B20D296 CRC64;

Query Match 26.8%; Score 433; DB 13; Length 371;
Best Local Similarity 37.1%; Pred. No. 1.3e-33;
Matches 101; Conservative 50; Mismatches 97; Indels 24; Gaps 10;

QY 20 EAPSOVLNFEEDYKIEIEVEEVGAGFVGVCKAKW--RAKDVAIK---QIESESERKAF 74
Db 31 ENSSLASAFVQIPFDDIRFVENCGGSGFVYRAHWVPQDKEVAVKLLKIDAEAE--- 86
QY 75 IVELRSLRVNHPNIVKLYGACLNVP--CLNMEYABGGSGLYNVLHGAEPLPYTTAAHAMS 132
Db 87 -----ILSVLSHKNIITQFYGAILEAFNDGIVTEYASRGSLEYLSADS--EEMDMQVMT 140
QY 133 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 190
Db 141 WAMEIAKGHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASKWVHTTHMSLV 198
QY 191 GSAAMMAPEVFEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRP 249
Db 199 GTFPMAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFK--GFEGQLQVAVLWVEKHERP 256
QY 250 PLIKNPKPIESLMTKRCWSDPSQSPSMEIEIV 281
Db 257 TIPSSCPASFAELMRRRCWNAEPKRPQFKQIL 288

RESULT 15
Q9HCC4 PRELIMINARY; PRT; 455 AA.
AC Q9HCC4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
```

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE MLTK-beta (Similar to sterile-alpha motif and leucine zipper
DE containing kinase AZK) (Mixed lineage kinase) (Mixed lineage
DE kinase-related kinase MRK-beta).
GN MLTK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1].
RP SEQUENCE FROM N.A.
RA Gotoh I., Adachi M., Nishida E.;
RT "Identification and Characterization of a Novel MAP Kinase Kinase
RT Kinase, MLTK.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [2].
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [3].
RP SEQUENCE FROM N.A.
RA Acton S.;
RT "MLK-mixed lineage kinase.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [4].
RP SEQUENCE FROM N.A.
RX MEDLINE=21950776; PubMed=11836244;
RA Gross E.A., Callow M.G., Waldbaum L., Thomas S., Ruggieri R.;
RT "MRK, a Mixed Lineage Kinase-related Molecule That Plays a Role in
RT gamma-Radiation-induced Cell Cycle Arrest.";
RL J. Biol. Chem. 277:13873-13882(2002).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB049734; BAB16445.1; -.
DR EMBL; BC001401; AAO01401.1; -.
DR EMBL; AF325454; AAK11615.1; -.
DR EMBL; AF480462; AAL85892.1; -.
DR HSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; Tyr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR SMART; SM00219; TYRK; 1.
DR SMART; SM00220; S_TK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 455 AA; 51582 MW; E87DB84A4D58B752 CRC64;
```

Query Match 26.8%; Score 433; DB 4; Length 455;

Best Local Similarity 36.7%; Pred. No. 1.7e-33;

Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

```
QY 27 NFEIDYKETEVEVGVGRGAFVGVCKAKW--RAKOVAIK---QIESESERKAFIVELRQL 81
Db 7 SFVQIKFDDQLQFENCGGSGFVYRAKMTSQDKEVAVKLLKIEKAE-----IL 57
QY 82 SRVNHNPVVKLYGACLNVP--VCLVMEYABGGSGLYNVLHG--AEPLPYTTAAHAMSCLQC 137
Db 58 SVLSHRNIITQFYGVILEPPNYGIVTEYASLGSGLYDINNRSEEM---DMDHIMTWATDV 114
QY 138 SOGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNK--GSAAW 195
Db 115 AKGMHYLHWEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRFHNHTTHMSLVGTFPW 172
QY 196 MAPEVFEGSNYSEKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
Db 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVPFKGLEG--LQVAVLWVEKNERLITPSS 230
```



```
Qy      255 LPKPIESIMTRCNSKIDPSQRPSPMEIVKIM 284
      | : : : | : : : | : : : | : : : |
Db      231 CPRSFAELLHQCEADAKKRPSPFKQIISIL 260
```

Search completed: December 10, 2002, 03:50:38
Job time : 98 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:46:30 ; Search time 38 Seconds
(without alignments)
234.609 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303
Perfect score: 1615
Sequence: 1 MSTASASSSSSSSAGEMIE.....MTHLMRYFPQADEPLQYPCQ 303

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued Patents_AA.*
- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	579	4	US-09-529-279-4
2	1615	100.0	590	4	US-09-529-279-15
3	433	26.8	455	3	US-09-221-235-5
4	433	26.8	455	3	US-09-221-928-5
5	433	26.8	455	4	US-09-221-527-5
6	433	26.8	455	4	US-09-221-236-5
7	433	26.8	455	4	US-09-221-416-5
8	433	26.8	455	4	US-09-221-245-5
9	433	26.8	455	4	US-09-163-115-5
10	433	26.8	455	4	US-09-221-528-5
11	433	26.8	455	4	US-09-593-553-5
12	433	26.8	455	4	US-09-221-237-5
13	410	25.4	835	4	US-09-291-839-2
14	401	24.8	668	1	US-08-205-018-2
15	401	24.8	859	1	US-08-395-580-2
16	401	24.8	859	5	PCT-US95-02792-2
17	398	24.6	821	1	US-07-928-464-2
18	398	24.6	821	1	US-08-003-311B-2
19	398	24.6	821	1	US-08-261-432-2
20	398	24.6	821	5	PCT-US93-07347-2
21	393	24.3	263	3	US-09-035-706-5
22	393	24.3	263	3	US-08-955-841-5
23	393	24.3	263	4	US-09-390-425-5
24	393	24.3	263	4	US-09-566-906-5
25	376.5	23.3	1584	4	US-09-457-040B-27
26	363.5	22.5	527	4	US-08-426-509A-10
27	363.5	22.5	527	5	PCT-US95-05008-10

28	358.5	22.2	620	4	US-08-426-509A-9	Sequence 9, Appli
29	358.5	22.2	620	5	PCT-US95-05008-9	Sequence 9, Appli
30	356	22.0	625	1	US-08-391-615-3	Sequence 3, Appli
31	351.5	21.8	269	2	US-07-857-224B-79	Sequence 79, Appli
32	350	21.7	983	1	US-08-167-919A-10	Sequence 10, Appli
33	350	21.7	983	3	US-08-715-106-10	Sequence 10, Appli
34	349.5	21.6	304	2	US-08-701-191A-27	Sequence 27, Appli
35	349	21.6	983	2	US-08-449-645A-21	Sequence 21, Appli
36	349	21.6	983	2	US-08-702-367A-21	Sequence 21, Appli
37	349	21.6	983	5	PCT-US95-04681-21	Sequence 4, Appli
38	346.5	21.5	982	2	US-08-673-789-4	Sequence 4, Appli
39	345	21.4	276	2	US-07-857-224B-72	Sequence 72, Appli
40	344.5	21.3	983	1	US-08-162-809-16	Sequence 16, Appli
41	341.5	21.1	316	1	US-08-278-089A-16	Sequence 16, Appli
42	341.5	21.1	316	2	US-08-838-957A-15	Sequence 15, Appli
43	341.5	21.1	588	1	US-08-391-615-5	Sequence 5, Appli
44	339.5	21.0	787	4	US-09-188-930-334	Sequence 334, App
45	338.5	21.0	261	2	US-07-857-224B-58	Sequence 58, Appli

ALIGNMENTS

RESULT 1
US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match	100.0%;	Score 1615;	DB 4;	Length 579;
Best Local Similarity	100.0%;	Pred. No. 5.2e-166;		
Matches 303;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MSTASASSSSSSSAGEMIEAPSOVLNFEELDYKEIEVEEVWGRGAFVYCKAKWRAKDV	60	
Db	1	MSTASASSSSSSSAGEMIEAPSOVLNFEELDYKEIEVEEVWGRGAFVYCKAKWRAKDV	60	
Qy	61	AIKQIESERKAFIVELRQLSRVNHENIVKLYGACLNPCVCLVMEYAEAGSLYNVLHGAE	120	
Db	61	AIKQIESERKAFIVELRQLSRVNHENIVKLYGACLNPCVCLVMEYAEAGSLYNVLHGAE	120	
Qy	121	PLPYTTAAHANSWCLQCSQGVAYLHSMQPKALHRDLKPPNLLLVAGTGLKICDFGTAC	180	
Db	121	PLPYTTAAHANSWCLQCSQGVAYLHSMQPKALHRDLKPPNLLLVAGTGLKICDFGTAC	180	
Qy	181	DIOTHTMNTKNGSAAMPAVEPEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM	240	
Db	181	DIOTHTMNTKNGSAAMPAVEPEGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIM	240	
Qy	241	WAVHNGTRPPLIKNLKPPIESLMTKRCWKDPSQRPSEIEIVKIMTHLMRYFPFGADEPQY	300	
Db	241	WAVHNGTRPPLIKNLKPPIESLMTKRCWKDPSQRPSEIEIVKIMTHLMRYFPFGADEPQY	300	
Qy	301	PCQ 303		

```
Db      301 PCQ 303

RESULT 2
US-09-529-279-15
; Sequence 15, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 590
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-15

Query Match      100.0%; Score 1615; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 5,3e-166;
Matches 303; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MSTASASSSSSSAGEMEAPSOVINEEIDYKEIEVEEVVGRGAFGVCAKRRADV 60
Db      1 MSTASASSSSSSAGEMEAPSOVINEEIDYKEIEVEEVVGRGAFGVCAKRRADV 60
QY      61 AIKQIESESEKAFIYELRQLSRVNHPNIVKLYGACINPCLVMEYAEGLSYNVLHGAE 120
Db      61 AIKQIESESEKAFIYELRQLSRVNHPNIVKLYGACINPCLVMEYAEGLSYNVLHGAE 120
QY      121 PLPYRYTAHAHMSWCLQCSGVAYLHSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
Db      121 PLPYRYTAHAHMSWCLQCSGVAYLHSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTAC 180
QY      181 DIQHTMTNKGSAAMWAPVEFGSNVSEKCDVFSWGIILMEVITRRKPFDEIGAPFRIM 240
Db      181 DIQHTMTNKGSAAMWAPVEFGSNVSEKCDVFSWGIILMEVITRRKPFDEIGAPFRIM 240
QY      241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMEIIVKIMTHLMRYFGADEPLQY 300
Db      241 WAVHNGTRPPLIKNLPKPIESLMTRCWSKDPSPQSPMEIIVKIMTHLMRYFGADEPLQY 300
QY      301 PCQ 303
Db      301 PCQ 303

RESULT 3
US-09-221-235-5
; Sequence 5, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-221-235-5

Query Match      26.8%; Score 433; DB 3; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY      27 NFEIDYKEIEVEEVVGRGAFGVCAKRW--RAKDVAK--QIESESEKAFIYELROL 81
Db      7 SFVQIKFDLQFFENCGGSGFSYVRAKWIQDKEVAVAKLLIKIEKAE-----IL 57
QY      82 SRVNHPIVLYGACINP--VCLVMEYAEGLSYNVLHG--AEPPLRYTAHAHMSWCLQC 137
Db      58 SVLSHRNIIQFYGVILPEPNYGVITEYASLGSLYDINSNRSEEM---DMDHIMTWATDV 114
QY      138 SQGVAYLHSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQHTMTNKK--GSAAM 195
Db      115 AKGMHYLHMEAPYKVIHRDLKSRNVVIAADG--VLKICDFG--ASRPHNHTTHMSLVGTFFW 172
QY      196 MAPEVEFGSNVSEKCDVFSWGIILMEVITRRKPFDEIGAPFRIMW-AVHNGTRPPLIKN 254
Db      173 MAPEVIOQLPVESETCDTYSYGVVLMEMLTREVPFKLEG--LQVAMLVYKNERLTIPS 230
QY      255 LKPIESLMTRCWSKDPSPQSPMEIIVKIM 284
Db      231 CPRSFALLHQCEWADAKKRPSFKQIISIL 260

RESULT 4
US-09-221-928-5
; Sequence 5, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-5

Query Match      26.8%; Score 433; DB 3; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY      27 NFEIDYKEIEVEEVVGRGAFGVCAKRW--RAKDVAK--QIESESEKAFIYELROL 81
Db      7 SFVQIKFDLQFFENCGGSGFSYVRAKWIQDKEVAVAKLLIKIEKAE-----IL 57
QY      82 SRVNHPIVLYGACINP--VCLVMEYAEGLSYNVLHG--AEPPLRYTAHAHMSWCLQC 137
Db      58 SVLSHRNIIQFYGVILPEPNYGVITEYASLGSLYDINSNRSEEM---DMDHIMTWATDV 114
QY      138 SQGVAYLHSMOPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQHTMTNKK--GSAAM 195
Db      115 AKGMHYLHMEAPYKVIHRDLKSRNVVIAADG--VLKICDFG--ASRPHNHTTHMSLVGTFFW 172
QY      196 MAPEVEFGSNVSEKCDVFSWGIILMEVITRRKPFDEIGAPFRIMW-AVHNGTRPPLIKN 254
Db      173 MAPEVIOQLPVESETCDTYSYGVVLMEMLTREVPFKLEG--LQVAMLVYKNERLTIPS 230
QY      255 LKPIESLMTRCWSKDPSPQSPMEIIVKIM 284
Db      231 CPRSFALLHQCEWADAKKRPSFKQIISIL 260
```

```
RESULT 5
US-09-221-527-5
; Sequence 5, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-527-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57
```

```
QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

RESULT 7
US-09-221-416-5
; Sequence 5, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57

QY 82 SRVNHNPVVKLYGACLN--VCLWMEYAEAGGSLYNVLHG--AEPLPYTTAAHAMSWCLOC 137
DB 58 SVLSHRNIIQFYGVILEPPNYGIIVTEYASGLSDYINSRSEEM---DMDHIMTWATDV 114

QY 138 SOGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGTACDIQTHMTNNK--GSAAW 195
DB 115 AKGMHYLHMEAPVKVIHRLDKSRNVVIAADG-VLKICDFG-ASRPHNHTHMSLVGTFPW 172

QY 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMW-AVHNGTRPPLIKN 254
DB 173 MAPEVIQSLPVSETCDTYSYGVVLWEMLTREVVPFKLEG--LQVAVLVVEKNERLTIPSS 230

QY 255 LPKPISLMTRCWSKDPSPRSMEEIVKIM 284
DB 231 CPRSFAELLHQCEADAKKRPFSKQIISIL 260

; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-416-5

Query Match      26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2.3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

QY 27 NFEIDYKEIEVEVGRGAFGVCKAKW--RAKDVAIK---QIESESEKAFIVELRQL 81
DB 7 SFVQIKFDDLOFFENCGGGSGSVYRAKWIISQDKEVAVKLLKIEKEAE-----IL 57
```

```

; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-245-5

```

```

Query Match          26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

```

```

QY 27 NPEEDYKEIEVEEVVGRGAFGVCKAKW--RAKDVAIK---QIESESERKAFIVELROL 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 SFQIKFDDLOFPFENGCGGSGFSGVYRAKWIISODKEVAVKKLTIKEKEA-----IL 57

QY 82 SRVNHPIVYKLYGACLNLP--VCLVMEYABEGSLYNVLHG--AEPPLPYTAAHMSWCLOC 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEM--DMDHIMTWATDV 114

QY 138 SOGVAAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNNK--GSAAM 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 AKGMHYLHMEAPYKVIVHRDLKSRNVVIAADG-VLKICDPG-ASRFNNHTTHMSLVGTFFW 172

QY 196 MAPEVEGGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 MAPEVIOSLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAVLVYKNERLTIIPS 230

QY 255 LKPRIESLMTRCWSKDPSPQPSMEELIVKIM 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 CPRSFAELHQCWEADAKKRPSFKQIISIL 260

```

```

RESULT 9
US-09-163-115-5
; Sequence 5, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-163-115-5

```

```

Query Match          26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

```

```

QY 27 NPEEDYKEIEVEEVVGRGAFGVCKAKW--RAKDVAIK---QIESESERKAFIVELROL 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 SFQIKFDDLOFPFENGCGGSGFSGVYRAKWIISODKEVAVKKLTIKEKEA-----IL 57

QY 82 SRVNHPIVYKLYGACLNLP--VCLVMEYABEGSLYNVLHG--AEPPLPYTAAHMSWCLOC 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEM--DMDHIMTWATDV 114

QY 138 SOGVAAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNNK--GSAAM 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 AKGMHYLHMEAPYKVIVHRDLKSRNVVIAADG-VLKICDPG-ASRFNNHTTHMSLVGTFFW 172

QY 196 MAPEVEGGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 MAPEVIOSLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAVLVYKNERLTIIPS 230

QY 255 LKPRIESLMTRCWSKDPSPQPSMEELIVKIM 284

```

```

Db 231 CPRSFAELHQCWEADAKKRPSFKQIISIL 260

```

```

RESULT 10
US-09-221-528-5
; Sequence 5, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-528-5

```

```

Query Match          26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

```

```

QY 27 NPEEDYKEIEVEEVVGRGAFGVCKAKW--RAKDVAIK---QIESESERKAFIVELROL 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 7 SFQIKFDDLOFPFENGCGGSGFSGVYRAKWIISODKEVAVKKLTIKEKEA-----IL 57

QY 82 SRVNHPIVYKLYGACLNLP--VCLVMEYABEGSLYNVLHG--AEPPLPYTAAHMSWCLOC 137
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 58 SVLSHRNIIQFYGVILEPPNYGIVTEYASLSGLYDINSNRSEEM--DMDHIMTWATDV 114

QY 138 SOGVAAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTACDIQTHMTNNK--GSAAM 195
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 115 AKGMHYLHMEAPYKVIVHRDLKSRNVVIAADG-VLKICDPG-ASRFNNHTTHMSLVGTFFW 172

QY 196 MAPEVEGGSNYSEKCDVFSWGIIIMEVITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 173 MAPEVIOSLPVSECTDYSGVGLWEMLTREVPFKGLEG--LQVAVLVYKNERLTIIPS 230

QY 255 LKPRIESLMTRCWSKDPSPQPSMEELIVKIM 284
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 231 CPRSFAELHQCWEADAKKRPSFKQIISIL 260

```

```

RESULT 11
US-09-593-553-5
; Sequence 5, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMT-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-593-553-5

```

```

Query Match          26.8%; Score 433; DB 4; Length 455;
Best Local Similarity 36.7%; Pred. No. 2,3e-38;
Matches 99; Conservative 53; Mismatches 90; Indels 28; Gaps 11;

```

Qy	27	NFBEIDYKEI	TEVEV	VG	RG	FG	VG	VCK	AKW	--RAK	VAIK	--QIE	SE	SER	KAFI	VEL	ROL	81																														
Db	7	SFVQIKEDDI	QF	FENC	GGG	FG	SV	RY	AKR	YSQ	DE	VAV	KLL	KIE	BAE	-----IL	57																															
Qy	82	SRVNHPI	NIV	KLY	GAC	LNP	--V	CL	VM	EY	AE	GG	S	LY	N	LHG	--AB	PL	PY	TA	AA	HA	MS	CL	OC	137																						
Db	58	SVLSHRNI	I	QF	Y	G	V	ILE	PP	NY	G	IV	E	Y	AS	L	S	LY	D	Y	NS	NR	SE	---	D	WD	H	IM	T	WA	T	VD	114															
Qy	138	SOQVAYI	LH	SQ	P	K	AL	I	HR	DL	K	P	N	L	L	V	AG	T	V	L	K	C	D	F	G	T	A	C	D	I	O	T	H	M	T	N	N	K	---	G	SA	A	195					
Db	115	AKGMHYI	LHE	AP	KV	I	HR	DL	K	R	N	V	I	A	A	D	G	--V	L	K	C	D	F	G	--A	S	R	F	N	H	T	H	T	H	S	L	V	G	T	F	F	W	172					
Qy	196	MAPEVFG	S	GS	Y	N	SE	K	D	V	F	S	G	I	L	W	E	V	I	T	R	R	K	P	D	E	I	G	G	P	A	P	I	M	--A	V	H	N	G	T	R	P	P	L	I	N	K	254
Db	173	MAPEVIO	S	L	P	S	E	T	D	I	Y	S	G	V	L	W	E	M	T	R	E	V	F	P	K	G	L	E	G	--L	Q	V	A	M	L	V	E	K	N	E	R	L	T	I	P	S	230	
Qy	255	LPKPIES	L	M	T	R	C	W	S	K	D	P	S	R	P	S	E	E	I	V	K	M	284																									
Db	231	CPRSFA	ELL	H	O	C	E	A	D	A	K	P	S	P	K	O	I	I	S	L	260																											

```

RESULT 12
US-09-221-237-5
; Sequence 5, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 455
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-237-5

```

Query Match	26.8%;	Score 433;	DB 4;	Length 455;
Best Local Similarity	36.7%;	Pred. No. 2.3e-38;		
Matches	99;	Conservative 53;	Mismatches 90;	Indels 28; Gaps
Qy	27	NFEISDYKEITEVEVWGEGAFGVVCKAKW--RAKDVAIK--QIESSEKKAIVELROL 81		
Db	7	SFVQIKFDLQFFENCGGSGFSGYRAKWSQDKEVAVKKLLKIEKAE-----IL 57		
Qy	82	SRVNHPIVTKLYGACLNP--VCLWEVYAEGLSLXNVILHG--ASPLPYTTAAHAMSCLQC 137		
Db	58	SVLSHRNIIQYGVILPEPNYGIITEVASGLSLDYINSNRSEM---DMDHIMTWATDV 114		
Qy	138	SGQVAYLHSMOPKALIHRLDLPPNLLVAGTGLVKICDFGTACDIOIHTMTNNK--GSAAW 195		
Db	115	AKGHHYLHWEAPVKVIHRDLKSRNVVIAADG-VLKICDFG-ASRFNNHTTHMSLVGTFPW 172		
Qy	196	MAPEVFEGSNYSEKCDVFSWGIIILWEITRRKPFDEIGGPAFRIMW-AVHNGTRPPLIKN 254		
Db	173	MAPEVIQSLPVSETCDTYSYGWLWEMLTREVPPKGLG--LQVAWLIVVEKNERLTIPSS 230		
Qy	255	LPKPIESLMTRCWKDQSPQSPMEIEVIKIM 284		
Db	231	CPRSFAELLHQCWEADAKRFSFQIISIL 260		

RESULT 13
US-09-291-839-2
; Sequence 2, Application US/09291839A
; Patent No. 6261818
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

LENGTH: 668 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-205-018-2

Query Match 24.8%; Score 401; DB 1; Length 668;
Best Local Similarity 33.5%; Pred. No. 1,2e-34;
Matches 88; Conservative 54; Mismatches 99; Indels 22; Gaps 9;

QY 30 EIDYKEIEVEEYVGRGAFVVCCKAKWRADVAIKQIESSEERKAFIVELRQLSRVNHNI 89
DB 119 EVPEEILDLQWVGSGAQAVFLGRPHGEVAVKKYRDKE-----TDIKHLRKLGHNPI 173
QY 90 VKLYGACLP--VCLVMEYAEAGSLYNNVLHGAEPLPYTTAAHMSWCLOCSQGVAYLHSM 147
DB 174 ITPKGVCTQAPCCYCIIMEFCAQGLYEVLRAGRPV---TPSLVDMSWGIAQNMNYLHLH 230
QY 148 QPKALIHRLDKPNNLLVAGTVLKI CDFTACDIQTHMTNNK--GSAAMMAPEVFECSN 205
DB 231 K--IIHRDLKSPN-MLITYDDVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
QY 206 YSEKCDVFSMGITLMEVITRRKPFDEIGAPAFRIMAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVVLWELLTGEIPIYKDVSAA--IIMGVGSNSLHLPVSSCPDGFKILR 344
QY 265 RCWSKDPQSQRPSMEIYKIMTHL 287
DB 345 QCWNSKPRNRPSPRQ---ILLHL 364

RESULT 15
US-08-395-580-2
Sequence 2, Application US/08395580
Patent No. 5676945

GENERAL INFORMATION:

APPLICANT: Usharani R. Reddy, David Pleasure and the Children's

TITLE OF INVENTION: Hospital of Philadelphia

TITLE OF INVENTION: No. 5676945el Protein Kinase, Nucleic Acid

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESS: Woodcock Washburn Kurtz Mackiewicz and No. 5676945trls

STREET: One Liberty Place - 46th Floor

CITY: Philadelphia

STATE: PA

COUNTRY: U.S.A.

ZIP: 19103

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch disk, 720 KB

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Wordperfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/395,580

FILING DATE: herewith

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/205,018

FILING DATE: 01-MAR-1994

ATTORNEY/AGENT INFORMATION:

NAME: Rebecca L. Ralph (formerly Gaumond)

REGISTRATION NUMBER: 35,152

REFERENCE/DOCKET NUMBER: CH-0488

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100

TELEFAX: 215-568-3439

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 859 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-395-580-2

Query Match 24.8%; Score 401; DB 1; Length 859;
Best Local Similarity 33.5%; Pred. No. 1,7e-34;
Matches 88; Conservative 54; Mismatches 99; Indels 22; Gaps 9;

QY 30 EIDYKEIEVEEYVGRGAFVVCCKAKWRADVAIKQIESSEERKAFIVELRQLSRVNHNI 89
DB 119 EVPEEILDLQWVGSGAQAVFLGRPHGEVAVKKYRDKE-----TDIKHLRKLGHNPI 173
QY 90 VKLYGACLP--VCLVMEYAEAGSLYNNVLHGAEPLPYTTAAHMSWCLOCSQGVAYLHSM 147
DB 174 ITPKGVCTQAPCCYCIIMEFCAQGLYEVLRAGRPV---TPSLVDMSWGIAQNMNYLHLH 230
QY 148 QPKALIHRLDKPNNLLVAGTVLKI CDFTACDIQTHMTNNK--GSAAMMAPEVFECSN 205
DB 231 K--IIHRDLKSPN-MLITYDDVKISDFGTSKELSDKSTKMSFAGTVAMMAPEVIRNEP 286
QY 206 YSEKCDVFSMGITLMEVITRRKPFDEIGAPAFRIMAV-HNGTRPPLIKNLPKPIESLMT 264
DB 287 VSEKVDIWSFGVVLWELLTGEIPIYKDVSAA--IIMGVGSNSLHLPVSSCPDGFKILR 344
QY 265 RCWSKDPQSQRPSMEIYKIMTHL 287
DB 345 QCWNSKPRNRPSPRQ---ILLHL 364

Search completed: December 10, 2002, 03:52:38
Job time : 42 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:48:00 ; Search time 214 Seconds
(without alignments)
22,997 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASASSSSSSAGEMIE.....MTHLMRYPPGADEPLQVPCQ 303

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW PUB.pap.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT NEW PUB.pap.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW PUB.pap.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pap.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW PUB.pap.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pap.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pap.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pap.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW PUB.pap.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pap.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW PUB.pap.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pap.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW PUB.pap.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	579	9	US-10-158-895-4
2	1615	100.0	590	9	US-10-158-895-15
3	433	26.8	455	10	US-09-757-982-5
4	432	26.7	1036	12	US-10-014-882-2
5	418.5	25.9	394	10	US-09-862-027-19
6	416	25.8	328	10	US-09-862-027-18
7	410	25.4	835	10	US-09-947-199-2
8	410	25.4	835	10	US-09-947-199-8
9	399	24.7	966	10	US-09-771-161A-197
10	393	24.3	263	10	US-09-840-704-5
11	392	24.3	850	10	US-09-904-389-2
12	369	22.8	425	10	US-09-828-313-29
13	363.5	22.5	527	10	US-09-977-269-10
14	360.5	22.3	277	10	US-09-882-166-4
15	359	22.2	278	9	US-09-842-582-4
16	359	22.2	278	10	US-09-797-039-13
17	359	22.2	278	10	US-09-922-138-18
18	359	22.2	278	10	US-09-922-138-27
19	359	22.2	278	10	US-09-910-150-17

20	359	22.2	278	10	US-09-910-150-31	Sequence 31, Appl
21	358.5	22.2	277	10	US-09-815-915-13	Sequence 13, Appl
22	358.5	22.2	620	10	US-09-977-269-9	Sequence 9, Appl
23	355	22.0	278	10	US-09-815-915-16	Sequence 16, Appl
24	351	21.7	273	10	US-09-922-138-11	Sequence 11, Appl
25	350	21.7	271	10	US-09-799-875-28	Sequence 28, Appl
26	350	21.7	272	10	US-09-780-949-4	Sequence 4, Appl
27	350	21.7	272	10	US-09-910-150-27	Sequence 27, Appl
28	349	21.6	983	10	US-09-771-161A-227	Sequence 23, Appl
29	348.5	21.6	279	10	US-09-799-875-23	Sequence 23, Appl
30	343	21.2	272	10	US-09-815-915-14	Sequence 14, Appl
31	333	20.6	507	10	US-09-977-269-2	Sequence 2, Appl
32	331.5	20.5	280	10	US-09-515-806-15	Sequence 15, Appl
33	329.5	20.4	928	8	US-08-578-684-2	Sequence 2, Appl
34	329	20.4	265	10	US-09-797-039-10	Sequence 10, Appl
35	329	20.4	675	9	US-10-186-399-3	Sequence 3, Appl
36	329	20.4	675	10	US-09-977-269-4	Sequence 4, Appl
37	329	20.4	764	10	US-09-925-302-714	Sequence 714, App
38	323.5	20.0	271	10	US-09-840-704-6	Sequence 6, Appl
39	320	19.8	1052	10	US-09-757-100B-2	Sequence 2, Appl
40	319.5	19.8	1276	10	US-09-982-610-24	Sequence 24, Appl
41	318.5	19.7	239	10	US-09-797-039-12	Sequence 12, Appl
42	318.5	19.7	245	10	US-09-815-915-15	Sequence 15, Appl
43	318.5	19.7	628	10	US-09-862-027-48	Sequence 48, Appl
44	318.5	19.7	1104	10	US-09-982-610-36	Sequence 36, Appl
45	317.5	19.7	212	10	US-09-834-496A-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 033466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match	100.0%;	Score	1615;	DB	9;	Length	579;
Best Local Similarity	100.0%;	Pred. No.	2.1e-124;				
Matches	303;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKD	60				
Db	1	MSTASASSSSSSAGEMIEAPSOVLNFEEDYKEIEVEEVVGRGAFGVCKAKWRAKD	60				
Qy	61	AIKQIESESRKAFIVELRQLSRVNHNPVVKLYGACLNPNVCLVMEYAEAGGSYLVNLHGA	120				
Db	61	AIKQIESESRKAFIVELRQLSRVNHNPVVKLYGACLNPNVCLVMEYAEAGGSYLVNLHGA	120				
Qy	121	PLPYTTAAHAMSWCLQCSQGVAYVYHSMOPKALIHRLDKPPNLLLVAGTGVLIKICDFGTAC	180				
Db	121	PLPYTTAAHAMSWCLQCSQGVAYVYHSMOPKALIHRLDKPPNLLLVAGTGVLIKICDFGTAC	180				

```

OY      181  DIQMTMNNKSSAAMAAEVEEGSNYSKCVFSWGJILMVEITRRKPFDEIGGPARIM 240
DB      181  DIQMTMNNKSSAAMAAEVEEGSNYSKCVFSWGJILMVEITRRKPFDEIGGPARIM 240
OY      241  WAWHNGTRPPLIKLKPRIEISLMTRCWSKDPORSMEIYKINTHLMRYFPGADEPLQY 300
DB      241  WAWHNGTRPPLIKLKPRIEISLMTRCWSKDPORSMEIYKINTHLMRYFPGADEPLQY 300
OY      301  PCQ 303
DB      301  PCQ 303

RESULT 2
US-10-158-895-15
Sequence 15, Application US/10158895
Patent No. US2002015562A1
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIOKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/10/158, 895
CURRENT FILING DATE: 2002-06-03
PRIOR APPLICATION NUMBER: US/09/529, 279
PRIOR FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-32
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRT
ORGANISM: Homo sapiens
US-10-158-895-15

```

Query Match	100.0%	Score	1615;	DB	9;	Length	590;
Best Local Similarity	100.0%	Pred. No.	2	1e-124;			
Matches	303;	Conservative	0;	Mismatches	0;	Indels	0;
						Gaps	0;

QY	1	MSTASASSSSSSSAGEMIEAPSOVLNFEEDIDKEIEVEEVVGRGAFGVVCAKAKRAXDV	60
Db	1	MSTASASSSSSSSAGEMIEAPSOVLNFEEDIDKEIEVEEVVGRGAFGVVCAKAKRAXDV	60
QY	61	AIKQIESSSERAFIVELRQLSRVNHPTNIVKIQGACLNVCVLMEYAEBSGLYNNVLHGAE	120
Db	61	AIKQIESSSERAFIVELRQLSRVNHPTNIVKIQGACLNVCVLMEYAEBSGLYNNVLHGAE	120
QY	121	PLPYTTAAHAMSMWCQCSQGVAYLHSMOPKALIHFDLKPNNLLVAGGTVLKICDPGTFAC	180
Db	121	PLPYTTAAHAMSMWCQCSQGVAYLHSMOPKALIHFDLKPNNLLVAGGTVLKICDPGTFAC	180
QY	181	DIQTHMTNNKSGAAMWAPPEVFEGSYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM	240
Db	181	DIQTHMTNNKSGAAMWAPPEVFEGSYSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIM	240
QY	241	WAVNHGTRPPLIKNLPKPLESLMTGCSKSDPQRSMETIVKIMTHLMRYFFPGADEPLQY	300
Db	241	WAVNHGTRPPLIKNLPKPLESLMTGCSKSDPQRSMETIVKIMTHLMRYFFPGADEPLQY	300
QY	301	PCQ 303	
Db	301	PCQ 303	

RESULT	3
US-09-757-982-5	
; Sequence 5, Application US/09757982	
Patent No. US2002094559A1	
GENERAL INFORMATION:	

```

? APPLICANT: Acton, Susan
? TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
? FILE REFERENCE: NMT-050
? CURRENT APPLICATION NUMBER: US/09/757,982
? CURRENT FILING DATE: 2001-01-10
? PRIOR APPLICATION NUMBER: 09/163,115
? PRIOR FILING DATE: 1998-09-29
? NUMBER OF SEQ ID NOS: 15
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 5
? LENGTH: 455
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-757-982-5

```

Query Match	26.8%;	Score 433;	DB 10;	Length 455;
Best Local Similarity	36.7%;	Pred. No. 3.9e-28;		
Matches	99;	Conservative	53;	Mismatches 90; Indels 28; Gaps 11;

QY	27	NFEELDYKEIVEEVEVVGARGAVYCCAKW--RAVDVAIK--QIESSEKAFVLELQOL	81
Db	7	SFVQIKEDFDLFFENCCGGSGFSGVYRAKMIISQDEVAIVAKLLIKIEKE-----IL	57
QY	82	SRVHNPVTKLYGACLNPF--VCLIMEAEGSLXNVLHG--AEPPLYTAHAAMWCLOC	137
Db	58	SVLHRNIIQYGVILDEPPNYGIETVEASLSGLVDYINSNISEEM---DMDHMTWADV	114
QY	138	SQSAVYLIHSMQPKALIRHDLKPPMLLVAGSTVYKICDFGACIOQTHMTNNK--GSAAM	195
Db	115	AKGMHYLHMEAPVYVIRHDELSRNVVLAADQ-VKICDFG--ASRFHNHTTMSLVGTFPW	172
QY	196	MAPEVEFGSSNASEKQDYFWSGIIIMVEYITRRKPEDEIGSPAFRIIMV-AVHNQTRPPLIKN	254
Db	173	MAPEVIGSLPVSETCIDTNYSGVVLMEMLTIREVPFKGLEG--LQVAMLVVENKEBLLTIPSS	230
QY	255	LKPRIESLMTRCWSKDSQRPMSMEIYKIM	284
Db	231	CPRSFAEILHOCWEADAKKRPSPFKQIISTL	260

```

RESULT 4
US-10-014-882-2
? Sequence 2, Application US/10014882
? Patent No. US20020107384A1
? GENERAL INFORMATION:
? APPLICANT: Hu, Yi
? APPLICANT: Kieke, James
? APPLICANT: Donoho, Gregory
? TITLE OF INVENTION: No. US-20020107384A1e1 Human Kinase and Polynucleotides Encoding t
? FILE REFERENCE: LEX-0279-USA
? CURRENT APPLICATION NUMBER: US/10/014,882
? CURRENT FILING DATE: 2001-12-11
? PRIOR APPLICATION NUMBER: US 60/254,744
? PRIOR FILING DATE: 2000-12-11
? NUMBER OF SEQ ID NOS: 3
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 2
? LENGTH: 1036
? TYPE: PRT
? ORGANISM: homo sapiens
? FEATURE:
? NAME/KEY: VARIANT
? LOCATION: (1)...(1036)
? OTHER INFORMATION: Xaa = Any Amino Acid
? US-10-014-882-2

```

```

Query Match 26.7% Score 432 DB 12 Length 1036;
Best Local Similarity 34.0% Pred. No. 1.2e-77;
Matches 99; Conservative 57; Mismatches 107; Indels 29; Gaps 7.

22 PSQVLFEEIDYKELEVEBVGREGAFGVCAKARRADVAIKQISESEKATIV----- 76
      :::::
110 PSSPSSPVHAERELKELKELGAGFGGVVATATQOGEVAVKAROPEDDAAAASVSR 169
      :::::

```



```
Db 576 YLHNLTOP--IHRDLNSHILLYEDGHAV-VADFESEFLOSLEDNDNTKOPGILRMWA 632
Qy 198 PEVF-EGSNYSEKCDVFSWGIIIMEVITRRKPEDEIGAPRIMAVHNGTRPPLIKNLP 256
Db 633 PEVFTCTRTYTIKADVFSAALCIMEILTGELIPFAHLKPAADAAVMAYHH-IRPPIGYSIP 691
Qy 257 KPESIMTRCWSKDPQSPQSMEEIV 281
Db 692 KPISILLIRGMWACPEGRPEFSEV 716
```

RESULT 8

```
US-09-947-199-8
; Sequence 8, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 835
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-947-199-8
```

Query Match 25.4%; Score 410; DB 10; Length 835;

Best Local Similarity 37.0%; Pred. No. 5.7e-26; Matches 102; Conservative 50; Mismatches 100; Indels 24; Gaps 10;

```
Qy 20 EAPSOVLNFEELDYKEIEVEEVGARGVVCARKAKDAVAKIOE-----SESERKAF 74
Db 451 ELPSRF---HQLSEIEFHETIIGSGFGKYKGRCKRIYAIKRYANTCSQSDVMF 506
Qy 75 IVELQOLSRVNHPIVYKLGACL--NPVCLVMEYAEAGSLYVNHGAEPLPYTAAHAM 131
Db 507 CREVSIILQOLNHPCVQVQFVAGCLDPSQFAITQYISGSLFSLH--EOKRIIDLQSKL 564
Qy 132 SWCLQCSQGVAYLHSM-OPKALIHRLDKPNNLLVAGGVLTICDPGTACDIQT---HM 186
Db 565 IIAVVAKAMEYLHSLTOP--IHRDLNSHILLYEDGHAV-VADFESEFLOSLEDNDM 621
Qy 187 TNKKSAAAMAEVFE-EGSNYSEKCDVFSWGIIIMEVITRRKPEDEIGAPRIMAVHN 245
Db 622 TKQPGNLRMWMAEVEFTQCTRTYTIKADVFSAALCIMEILTGELIPFAHLKPAADAAVMAYHH 681
Qy 246 GTRPPLIKNPKIESIMTRCWSKDPQSPQSMEEIV 281
Db 682 -IRPPIGYSIPKISSILLIRGMWACPEGRPEFSEV 716
```

RESULT 9

```
US-09-771-161A-197
; Sequence 197, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: Levine, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771,161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
```

```
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 197
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-197
```

Query Match 24.7%; Score 399; DB 10; Length 966;

Best Local Similarity 34.2%; Pred. No. 5.3e-25; Matches 90; Conservative 48; Mismatches 103; Indels 22; Gaps 9;

```
Qy 30 EIDYKEIEVEEVGARGVVCARKAKDAVAKIOESESERKAFIVELQOLSRVNHPI 89
Db 162 EVPEEISELQWISGAGVFLGKFAEVAIKVREONE-----TDIKHLKLGHPNI 216
Qy 90 VKLYAGCLNP--VCLVMEYAEAGSLYVNHGAEPLPYTAAHAMSWCLQCSQGVAYLHSM 147
Db 217 IAFRGVCTQAPCYCIIMEYCAHGOYLEVLAKGKI---TPRLIVDMSTGIASGMYLHLH 273
Qy 148 QPKALIHRLDKPNNLLVAGGVLTICDPGTACDIQTHMTNNK--GSAAMAPVEVEGSN 205
Db 274 K---IHRDLKSPN-VLVHTTDAVKISDFGTSKELSDKSTKMSFACTVAMMAPEVIRNP 329
Qy 206 YSEKCDVFSWGIIIMEVITRRKPEDEIGAPRIMAV-HNGTRPPLIKNPKIESIMT 264
Db 330 VSEKVDIWSFGVVLWELTGELIPYKQVDSAA--IIWGVSNLSLHPSTCPDGFIKLM 387
Qy 265 RCWSKDPQSPQSPQSMEEIVKIMTHL 287
Db 388 QTWQSKPRNRNPSFRQ---TLMHL 407
```

RESULT 10

```
US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dethat, Shoukat
; APPLICANT: Hannigan, Greg
; TITLE OF INVENTION: Intergrin-Linked Kinase and its Uses
; FILE REFERENCE: KIN-2CON
; CURRENT APPLICATION NUMBER: US/09/840,704
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/566,906
; PRIOR FILING DATE: 2000-05-09
; PRIOR APPLICATION NUMBER: US08/752,345
; PRIOR FILING DATE: 1996-11-19
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 263
; TYPE: PRT
; ORGANISM: H. sapiens
; FEATURE:
; NAME/KEY: Other
; LOCATION: (1)...(263)
US-09-840-704-5
```

Query Match 24.3%; Score 393; DB 10; Length 263;

Best Local Similarity 33.8%; Pred. No. 3.7e-25; Matches 90; Conservative 57; Mismatches 107; Indels 12; Gaps 8;

```
Qy 31 IDYKEIEVEEVGARGVVCARKAKDAVAKI--QIESER-KAFIVELQOLSRVNH 86
Db 1 IPWCDLINIKETIGAGSFVTHRAEWHGSDAVKIMEDQFAERVNEFLREVALIKRLRH 60
Qy 87 PNIVKLYGACLNLP--VCLVMEYAEAGSLYVNHGAEPLPYTAAHAMSWCLQCSQGVAYL 144
```

```
Db 61 PNIIVFMGAVTPQPNLSIVTEYLSRGLSYRLHLKSGAREQLDERRRLSMAYDVAKGNNYL 120
Qy 145 HSMQPKALIHRLDKPPNLLVAGGTVLKICDPG-TACDIQTHMTNNK--GSAAMWAEVVF 201
Db 121 HNRNP-PIVHRDLKSPNLLVDKKYTV-KVCDPGLSRLKASTFLSSKSAAGPFWMAPEVL 178
Qy 202 EGSNYSKCDVFSWGIILWEVITRRKPPDEIGGPAFRIMWAVHNGTTPPLIKNLPKPIES 261
Db 179 RDEPNEKSDVYSGFVILWELATLOQPNGL-NPAQVAAVGFCKRLLEIPRNLNPQAA 237
Qy 262 LMTRCWSKDPQSRPSMEIVKIMTHL 287
Db 238 IIEGCTWNEPWKRPSTFATIMDLLRL 263

RESULT 11
US-09-904-389-2
; Sequence 2, Application US/09904389
; Patent No. US20020129404A1
; GENERAL INFORMATION:
; APPLICANT: Clendennen, Stephanie K.
; APPLICANT: Schuster, Debra K.
; TITLE OF INVENTION: CTR1 HOMOLOGUE FROM MELON
; FILE REFERENCE: 4257-0029.30
; CURRENT APPLICATION NUMBER: US/09/904,389
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: US 60/218,307
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Cucumis melo
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (154)...(154)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-09-904-389-2

Query Match 24.3%; Score 392; DB 10; Length 850;
Best Local Similarity 32.1%; Pred. No. 1.7e-24;
Matches 93; Conservative 66; Mismatches 111; Indels 20; Gaps 12;

Qy 17 EMIEAPSOV-LNFE--IDYKEIEVEEVGCGAFGVCKAKWRAKDVAIKQIESES---E 70
Db 555 QLIRKPNELSLGLELIVPTDLDREKIGAGSGFTVYRGWHSADVAVKILTEQDPHPE 614
Qy 71 R-KAFIVELRQLSRVNHPIVVKLYGACLP--VCLVMEYAEAGGSLYVNLH--GASPLPY 125
Db 615 RVNEFLUREVAIMKSLRHPNIVFMGAVTPPNLSIVTEYLSRGLSYRLHLKSGVKDI--- 671
Qy 126 TAAHAMSCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPG-TACDIQT 184
Db 672 DETFRINMAPDVAKGNNYLHRRDP-PIVHRDLKSPNLLVDKKYTV-KVCDPGLSRLKART 729
Qy 185 HMTNNK--GSAAMWAEVFEFGSNYSKCDVFSWGIILWEVITRRKPPDEIGGPAFRIMWA 242
Db 730 FLSSKSAAGTPFWMAPEVLRLDEPNEKSDVYSGFVILWELATLQQPWCNL-NPAQVAAV 788
Qy 243 VHNGTTPPLIKNLPKPIESIMTRCWSKDPQSRPSMEIEIVKIMTHLMRYFP 292
Db 789 GFGKRLDIPRDVNPKNLASLIVACWADEPWKRPSTFSSIMETLKPMTKQAP 838

RESULT 12
US-09-828-313-29
; Sequence 29, Application US/09828313
; Patent No. US20020059662A1
; GENERAL INFORMATION:
; APPLICANT: COSTA e SILVA, OSWALDO DA
; APPLICANT: BOHNERT, HANS J.
```

```
; APPLICANT: THIELEN, NOCHA VAN
; APPLICANT: CHEN, ROUYING
; APPLICANT: SARRIA-MILIAN, RODRIGO
; TITLE OF INVENTION: PROTEIN KINASE STRESS-RELATED PROTEINS AND METHODS OF
; TITLE OF INVENTION: USE IN PLANTS
; FILE REFERENCE: 16313-0032
; CURRENT APPLICATION NUMBER: US/09/828,313
; CURRENT FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/196,001
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Physcomitrella patens
US-09-828-313-29

Query Match 22.8%; Score 369; DB 10; Length 425;
Best Local Similarity 33.8%; Pred. No. 5.8e-23;
Matches 94; Conservative 51; Mismatches 101; Indels 32; Gaps 10;

Qy 21 APSQVLNFEIDYKEIEVEEVGCGAFGVCKAKWRAKDVAIKQIESERKAFIV--- 76
Db 131 APSEI-----ELDTSELIGKAGFGEIRKALWRGTPVAVKTRPSSLNDRMVIKDFQ 181
Qy 77 -ELRQLSRVNHPIVVKLYGACL--NPVCLVMEYAEAGGSLYVNLHGAEPPLPYTTAAHMSW 133
Db 182 HEVQLLVKVRHPIVQFLGAVTRQRLMLVTEFLAGDLHQLLSRN---PNLAPDRIVK 238
Qy 134 CLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPGTA--CDIQT-----HM 186
Db 239 ALDIARGMSYLHN-RSKPIIHRDLKPRN-IIVDEEHELKVGDFGLSKLIDVKLMDHVYKM 296
Qy 187 TNNKGSAAWMAPEVFEFGSNYSKCDVFSWGIILWEVITRRKPPDEIGGPAFRIMWAVHNG 246
Db 297 TGGTGSYRYMAPEVFEHQPYDKSDVDFSGMILYEMFEGVAPPED--KQAYDAATLVARD 354
Qy 247 TRPPLIK--NLPKPIESIMTRCWSKDPQSRPSMEIEIVK 282
Db 355 DKRPMPRAQTYPPQMKALIEDCWSPTYTPKPPFVEIVK 392

RESULT 13
US-09-977-269-10
; Sequence 10, Application US/09977269
; Patent No. US20020082037A1
; GENERAL INFORMATION:
; APPLICANT: ULLRICH, AXEL
; APPLICANT: GISHIZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,269
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-977-269-10

Query Match 22.5%; Score 363.5; DB 10; Length 527;
Best Local Similarity 34.3%; Pred. No. 2.1e-22;
Matches 92; Conservative 48; Mismatches 111; Indels 17; Gaps 9;

Qy 30 BIDYKEIEVEEVGCGAFGVCKAKWRAK-DVAIKQI-ESESEKAFIVELRQLSRVNH 87
Db 260 EINPSELTFMRELGSGLFVGVVRLGKWRQAQYKVAIKAREGAMCEEDFIEAKVMKLTHP 319
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 01:12:20 ; Search time 3085 Seconds
(without alignments)
2858.396 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTAASASSSSSSAGEMIE.....MTHLMRYFFGADEPLQYPCQ 303

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-O=/cn2_1/USPTO spoal/US09830144/runat_04122002_141510_2911/app_query.fasta_1.455
-DB=GenEmbl -QPT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CN1_1.3637 @runat_04122002_141510_2911 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WAEN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database :

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.on.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1615	100.0	1704	6	E38399 NF-Kappa B
2	1615	100.0	1705	9	AB009358 Homo sapi
3	1615	100.0	1745	9	AF218074 Homo sapi
4	1615	100.0	2757	9	BC017715 Homo sapi
5	1615	100.0	2769	6	AX377912 Sequence
6	1615	100.0	2769	9	AB009356 Homo sapi
7	1615	100.0	2785	6	E38397 NF-kappa B
8	1615	100.0	2850	9	AB009357 Homo sapi
9	1615	100.0	2866	6	E38398 NF-kappa B
10	1609	99.6	2443	10	MUSTAK1
11	1609	99.6	3107	10	BC006665 Mus muscu
12	1490.5	92.3	2812	5	U92030 Xenopus lae
13	1414	87.6	135147	2	AC114407 Mus muscu
14	818	50.7	3349	3	AF199466 Drosophil
15	818	50.7	3386	3	AF051953 Drosophil
16	689.5	42.7	32564	2	AC014558 Drosophil
17	689.5	42.7	164942	3	AC011758 Drosophil
C 18	689.5	42.7	302303	3	AE003571 Drosophil
C 19	482.5	29.9	3138	9	HSMSTM
20	482.5	29.9	3454	6	AX337846 Sequence
21	482.5	29.9	3454	9	HARNMLK2
22	454	28.1	3911	9	AF251442
23	442	27.4	2120	6	AR119790 Sequence
24	442	27.4	2120	6	AR126750 Sequence
25	442	27.4	2120	6	AR128910 Sequence
26	442	27.4	2120	6	AR130841 Sequence
27	442	27.4	2120	6	AR138886 Sequence
28	442	27.4	2120	6	AR138886 Sequence
29	437	27.1	8125	3	AF481923
30	435	26.9	1429	10	AB049732
31	435	26.9	3146	10	AB049731 Mus muscu
32	434.5	26.9	1267	8	AY133876 Arabidops
33	434.5	26.9	1798	8	AY046026 Arabidops
34	434	26.9	1370	9	AF325454 Homo sapi
35	434	26.9	1571	9	BC001401 Homo sapi
36	434	26.9	1631	8	AY085535 Arabidops
37	434	26.9	1795	8	AY125513 Arabidops
38	434	26.9	2251	9	AK056310 Homo sapi
39	434	26.9	2631	9	AB030034 Homo sapi
40	434	26.9	3767	9	AF251441 Homo sapi
41	434	26.9	3884	9	AF480461 Homo sapi
42	434	26.9	7195	9	AF480462 Homo sapi
43	433	26.8	1365	6	AR119791 Sequence
44	433	26.8	1365	6	AR126751 Sequence
45	433	26.8	1365	6	AR128911 Sequence

ALIGNMENTS

RESULT 1

E38399	E38399	1704 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	NP-kappa B activation inhibitory drug targeting TAK1 and method for				
DEFINITION	identifying the same.				
ACCESSION	E38399.1	GI:18626979			
VERSION	JP 2000197500-A/5.				
KEYWORDS	unidentified.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1704)				
AUTHORS	Sugita,N., Sakurai,H., Kageyama,N. and Haegawa,H.				
TITLE	NP-kappa B activation inhibitory drug targeting TAK1 and method for				
JOURNAL	identifying the same				
COMMENT	Patent: JP 2000197500-A 5 18-JUL-2000;				
	TANABE SEIRYAKU CO LTD				
	OS Unidentified				
	PN JP 2000197500-A/5				
	PD 18-JUL-2000				
	PF 04-FEB-1999 JP 1999026803				
	PR				
	PI NAOHISA SUGITA, HIROAKI SAKURAI, NORIKO KAGEYAMA, PI HIROSHI				
	HASEGAWA				
	PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99,PC				
	C12Q1/02,				
	PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,				
	PC C12R1.91,				
	PC C12N5/00,C12N15/00,(C12N15/00,C12R1.91)				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	FM Key	Location/Qualifiers			
	FT source	1..1704			
FEATURES		Location/Qualifiers			
source	1..1704	/organism="unidentified"			
	/db_xref="taxon:32644"				
BASE COUNT	511 a 381 c 401 g 411 t				
ORIGIN					
	Alignment Scores:				
	Pred. No.:	2,52e-149	Length:	1704	
	Score:	1615.00	Matches:	303	
	Percent Similarity:	100.00%	Conservative:	0	
	Best Local Similarity:	100.00%	Mismatches:	0	
	Query Match:	100.00%	Indels:	0	
DB:		Gaps:	0		
US-09-830-144-2_COPY_1_303 (1-303) x E38399 (1-1704)					
OY	1 MetSerThrAlaSerAlaAlaSerSerSerSerAlaGluMetIleGlu	20			
DB	1 ATGCTCAAGCCTTCGCGCTCTCTCTCTCTCTGCTTCGCGCGGAATATCGAA	60			
OY	21 AlaProSerGlnValIleuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu	40			
DB	61 GCCCTTCCAGGCTCCCACTTGAAGATGCACTCAAGAGATCGAGGTGAAAG	120			
OY	41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrrArgAlaLysAspVal	60			
DB	121 GTTGTGGAGAGAGAGCCTTTGGAGTGTGTTGGCAAGCTCAAGTGGAGCAAAAGATGTT	180			
OY	61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80			
DB	181 GCTATTAAACAAATAGAAAGATGAATCTGAAGAGAAAGCCTTATTGTAGAGCTCGGAG	240			
OY	81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100			
DB	241 TTATCCGCTGGAAACCATCTCAATATTGTTAAAGCTTTATGGAGCGCTGTAATCCAGTG	300			
OY	101 CysLeuValMetGluTrrAlaGluGlyLysIleTyrAsnValLeuHisGlyValGlu	120			
DB	301 TGTCTTGAGTGAAGAAATGCTGAAGGGGGCTTTTATATAATGTCGTCGATGTCGAA	360			

QY	121	Proleuprotin	Y	Th	Ala	Ala	His	Ala	Met	Set	Trp	Cys	Leu	Gln	Cys	Ser	Gln	140			
Db	361	CCATTGCCATTTAATAC	T	ACT	GCGCCACGAA	TAGTGGTGTATACAG	GTGTTCCCAAG	A	420												
QY	141	Val	Ala	Ala	Thr	Leu	His	Ser	Met	Gln	Pro	Val	Ala	Leu	Ile	His	Arg	Ser	Leu	Pro	160
Db	421	GGGCTTATCTTCA	G	AGATCA	ACCACCA	AGGCTTAATTC	ACAGGAA	CTTGAA	CCGAA	CTGAA	CCACCA	480									
QY	161	Asn	Leu	Leu	Leu	Val	Ala	Gly	Thr	Val	Leu	His	Ile	Cys	Asp	Phe	Gly	Thr	Ala	Cys	180
Db	481	AACTACCTGCTG	T	GTGCA	GGGGAG	CAGTCTTAA	AAATTTGTAT	TTTGTGT	GTAC	AGCTGT	540										
QY	181	Asp	Ile	Gln	Thr	His	Met	Thr	Asn	Asn	His	Gly	Ser	Ile	Ala	Ala	Trp	Met	Leu	Pro	200
Db	541	GACATTCACAC	CA	CAATAC	CAATAC	CAAGGGAG	GTGCTGTG	ATGAC	CTTGAA	GTGCAC	CTTGAA	GT	600								
QY	201	Phe	Gly	Gly	Ser	Asn	Trp	Ser	Gly	Alu	Cys	Asp	Val	Phe	Ser	Trp	Ile	Ile	Leu	Trp	220
Db	601	TTTGAAGTAG	TA	ATTAC	AGTAA	AAATGT	AGAC	CTCTT	CAC	CTGGGG	TAT	ATCT	780								
QY	221	Glu	Val	Ile	Thr	Arg	Arg	Leu	Pro	Phe	Asp	Glu	Ile	Gly	Gly	Pro	Ala	Phe	Arg	Ile	240
Db	661	GAGGTATAC	CCGCGT	CGGAAA	CCCTTGA	TGATGAT	TGGTGG	CCCA	CTTCCG	AAATCA	720										
QY	241	Thr	Ala	Val	His	Asn	Gly	Thr	Arg	Pro	Pro	Leu	Ile	Leu	Asn	Leu	Pro	Lys	Pro	Ile	260
Db	721	TGGCGTTC	TA	AAATGA	AGTAC	TGAC	CA	CACTG	AT	AAAAAT	TTTAC	CTA	AGCC	ATT	780						
QY	261	Set	Leu	Met	Thr	Arg	Cys	Trp	Ser	Lys	Asp	Pro	Ser	Gln	Arg	Pro	Ser	Met	Glu	Ile	280
Db	781	AGCTGATAC	CTCG	TGTTGG	CTAA	GATC	CTTCC	CAGG	CGCC	CTTCA	ATGAG	GA	840								
QY	281	Val	Ile	Ile	Met	Thr	His	Leu	Met	Arg	Thr	Pro	Gly	Ala	Asp	Glu	Pro	Leu	Gln	Trp	300
Db	841	GTA	AAATAT	ATAT	CACT	CTGAT	TGGG	TA	CTTCC	AGAG	CA	AGAT	GA	CCAT	TAC	AG	900				
QY	301	Pro	Cys	Gln	303																
Db	901	CC	TTG	T	C	A	G	909													
RESULT 2																					
LOCUS	AB009358																				
DEFINITION	AB009358																				
ACCESSION	AB009358																				
VERSION	AB009358.2																				
KEYWORDS	TAKIC; TGF-beta																				
SOURCE	Human sapiens																				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.																				
REFERENCE	1 (sites)																				
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.																				
TITLE	TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism																				
JOURNAL	Biophys. Res. Commun. 243 (2), 545-549 (1998)		</																		

Db	541	GACATTCAGACACATGACCAATAACAGGGGAGTGCTGCTTGGATGGCCACCTGAAGTT	600
Qy	201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyValIleLeuTyr	220
Db	601	TTTGAAGGTAGTAAATTACAGTGAATAATGTGACGCTCTTCAGCTGGGTATTATTCTTTGG	660
Qy	221	GluValIleThrArgArgLysPropPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	661	GAAGTGATAACGCGTCGGAAACCCCTTTGATGAGATTGGTGGCCAGCTTTCGGAATCATG	720
Qy	241	TyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	721	TGGGCTGTTTCATAATGGTACTCGACCACCTGATAAAAAATTTTACCTAAGCCCATTTGAG	780
Qy	261	SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	781	AGCCTGATGACTCGTTTGGTCTAAAGATCTTCCAGCGCCCTTCAATGGAGGAATTT	840
Qy	281	ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	841	GTGAAAATAATGACTCACITTCATGCGGTACTTTCAGGAGCAGATGAGCCATTACAGTAT	900
Qy	301	ProCysGln	303
Db	901	CCTTGTTCAG	909
RESULT	3		
AF218074			
LOCUS	AF218074		
DEFINITION	Homo sapiens TGF beta-activated kinase splice variant d (TAK1)-		
ACCESSION	AF218074		
VERSION	AF218074.1	GI:6746614	
KEYWORDS			
SOURCE	Homo sapiens.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 1745)		
AUTHORS	Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.		
TITLE	Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1		
JOURNAL	Biochim. Biophys. Acta 1517 (1), 46-52 (2000)		
MEDLINE	20568765		
PUBMED	11118615		
REFERENCE	2 (bases 1 to 1745)		
AUTHORS	Dempsey,C.E. and Guesdon,F.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-DEC-1999) Division of Molecular and Genetic Medicine, University of Sheffield, School of Medicine, Glossop Road, Sheffield S10 2JF, United Kingdom		
FEATURES	Location/Qualifiers		
source	1..1745		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
gene	1..1745		
	/gene="TAK1"		
CDS	7..1482		
	/gene="TAK1"		
	/note="TAK1d"		
	/codon_start=1		
	/product="TGF beta-activated kinase splice variant d"		
	/protein_id="AAF27652.1"		
	/db_xref="GI:6746615"		
	/translation="MSTASAASSSSSSAGEMIEAPSQVLNFEEDYKEIEVEEVVGR		
	GAFGVCKAKRAKDVAIKOIIESESEKAFIVELRQLSRVNHPIVKYLGACLNVPCL		
	NLLVAGGTVLKICDFGTACDIOTHTNNKGAAMWAPVEFGSNYSEKCDVFSWGII		
	LWEVITRRKPFDFIGGPAFRIMWAVHNGTRPPLIKNLKPIESLTRCWSKDPSPORS		
	MEIVKIMTHLMRYFPGADEPLYPCQYSDGQSNASATSGFMDIASNTNSKSDTN		
	MEQVPATNDTIKKLESKLJNKAKQOSESRLSLGASRGSSVESLPPTSEGRKMSADM		
	SETEARIAATTAYSKPKRHKRTASFENLIDVPEIVISGNQPRRRSIQDLTVTQTEP		
	GOVSSRSSPSPVRMITTSPTSEKPTRSHPWTPDDSTDTNGSDNSIPMAYLTLDHQLQ		
	ARTSCRTGPG"		
BASE COUNT	511 a	381 c	402 g 411 t
ORIGIN			
Alignment Scores:			
Pred. No.:	2.52e-149	Length:	1705
Score:	1615.00	Matches:	303
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0
US-09-830-144-2_COPY_1_303 (1-303) x AB009358 (1-1705)			
Qy	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu	20
Db	1	ATGCTACAGCCTCGCGCGCTCCTCCTCCTCCTCGTTCGCGCGGTGAGATGATCGAA	60
Qy	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu	40
Db	61	GCCCCCTCCACAGTCTCAACTTTGAAGATCGACTACAGGAGATCGAGGTGGAAGAG	120
Qy	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal	60
Db	121	GTTGTTGGAAGAGAGCCTTTGAGTTGTTTGCAACCTAAGTGGAGAGCAAAAGATGT	180
Qy	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	181	GCTATTAAACAAATAGAAAGTGAATCTGAGAGGAAGCGTTTATTGTAGAGCTTCGCGAG	240
Qy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100
Db	241	TTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATCGAGCCCTGTTGAATCCAGTG	300
Qy	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	301	TGTCCTGTGATGGAATATGCTGAAGGGGCTCTTATATATAATGTCGATGCTGCTGAA	360
Qy	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly	140
Db	361	CCATTGCCATATTATACGTGCTGCCACGAATGAGTTGGTGTATTACAGTGTTCCTCAAGGA	420
Qy	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160
Db	421	GTGCTTATCTTACAGCATGCAACCCAAACGCGCTAATTCACAGGACCTGGAACACCA	480
Qy	161	AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	481	AACCTACTGCTGTTGAGGGGGACAGTTCTAAAAATTTGTAATTTGGTACAGCGCTGT	540
Qy	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal	200

[illegible]

Oy	301	Procysgin 303
Db	907	CCTGTGTCAG 915
RESULT 4		
LOCUS	BC017715	
DEFINITION	Homo sapiens, mitogen-activated protein kinase kinase kinase 7,	
ACCESSION	BC017715	
VERSION	BC017715.1	GI:17389342
KEYWORDS	MGC.	
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens.	
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 2757)	
JOURNAL	Strausberg, R.	
	Direct Submission	
	Submitted (03-DEC-2001) National Institutes of Health, Mammalian	
	Gene Collection (MGC), Cancer Genomics Office, National Cancer	
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,	
	USA	
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov	
COMMENT	Contact: MGC help desk	
	Email: gcgabs-remail.nih.gov	
	Tissue Procurement: ATCC	
	CDNA Library Preparation: Life Technologies, Inc.	
	DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)	
	Center, Stanford University School of Medicine, Stanford, CA 94305	
	Web site: http://www.sngc.stanford.edu	
	Contact: (Dickson, Mark) mcd@psxll.stanford.edu	
	Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,	
	R. M.	
	Clone distribution: MGC clone distribution information can be found	
	through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov	
	Series: IRAK Plate: 22 Row: 1 Column: 18	
	This clone was selected for full length sequencing because it	
	passed the following selection criteria: matched mRNA gt: 4507360.	
FEATURES	Location/Qualifiers	
SOURCE	1..2757	
	/organism="Homo sapiens"	
	/db_xref="LocustID:6885"	
	/db_xref="taxon:9606"	
	/clone="MGC:21263 IMAGE:3906837"	
	/tissue_type="Uterus, leiomyosarcoma"	
	/clone_id="NHL_MGC_71"	
	/lab_host="DH10B"	
	/note="Vector: pCMV-Sport6"	
	198..1937	
	/codon_start=1	
	/product="mitogen-activated protein kinase kinase kinase	
	7"	
	/protein_id="AAH17715.1"	
	/db_xref="GI:17389343"	
	/translation="MSPTSAASSSSSSSAGEMTEAPSVLNFEIDYKIEVEEVRGR	
	GAVGVCAKRAKRAVDIAKOIIESESEKATIVLRSLRNNPIVLYGACLNPCCLC	
	VMYAEGBGLYNLVLAHAEPYPYTAAHMSWCQSGVAYLHSMOKALIHRLKPKPP	
	NLLVLVGGLVIKICDFTACDICTDHTMTNNKGSAMMAPEVFBSNSSEKDFESWGIIL	
	LMEVITRRKPFDEIFGPADPPLPYIMAVNHNGTBPPLIKNPKEISLMTNCWSKDPSQRSP	
	MEQVATNDITIKLESLKLNOAQOSESRLSIGARSGSVESLPSTSGKMSADM	
	SHETPTDDSDTDNNGSDNSIMAYLTLDHQQLAPLCNSENMAVFEQCKMAOEYVKKK	
	VQTEIALILQRKQELVAELDDQDKDOONTSLRVQEHKKLLDENKSLSTYYQQCKQOLE	
	VIQSQQKRKGTS"	
BASE COUNT	819 a	566 c 647 g 725 t
ORIGIN		
Alignment Scores:		
Aligned No.:	4..46e-149	Length: 2757

```
Score: 1615.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x BC017715 (1-2757)

Qy 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerSerAlaGlyGluMetIleGlu 20
Db 198 ATGTCATACAGCCTTCGCCGCTTCCTCCTCCTCGCTTCGCGCGTGAGATGATCGAA 257
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu 40
Db 258 GCCCCTTCCAGGTCCTCAACTTGAAGAGATCGACTACAAGGAGATCGAGTGGAGAG 317
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal 60
Db 318 GTTGTGGAAGAGAGCCCTTTCGAGTGTGTTGCAAAAGCTAAGTGGAGAGCAAAAGATGTT 377
Qy 61 AlaIleLysGlnIleGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 378 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAAGCGTTTATTCTAGAGCTTCGGCAG 437
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
Db 438 TTATCCCGGTGGAACCATCTTAATATTGTAAGCTTTATGAGAGCCTGCTTGAATCCAGTG 497
Qy 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
Db 498 TGTCTTGTGATGGAATATCTGCAAGGGGCTCTTTATATATAATGTGCTGCATGGTCTGAA 557
Qy 121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 558 CCATTGCCATATTACTGCTGCCACGCAATGAGTGTGTTTACAGTGTTCCTCCACGA 617
Qy 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 160
Db 618 GTGGCTTATCTTCACAGCATGCAACCAAGCGCTAAATTCACAGGAGCTGAAACCA 677
Qy 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 678 AACTTACTGCTGCTGTCAGGGGGCAGGTTCTAAAAATTTGTGATTTTGGTACAGCCTGT 737
Qy 181 AsnIleGluThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 738 GACATTCAGACACACATGACCAATAACAAGGGAGTGCTGTGATGGACCTCGAAGTT 797
Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 798 TTTGAAAGGTAGTAATTACAGTGAAAAATGTGACGCTCTTCAGCTGGGGTATTATTCTTTGG 857
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 858 GAAGTGATAACCGTCGGAACCCCTTTGATGAGATTGGTGGCCACAGCTTTCGGAATCATG 917
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 918 TGGGCTGTTTCATTAATGGTACTGACCACCACTGATAAAAAATTTACCTAAGCCCATTTGAG 977
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 978 AGCTGATGACTGTTGTTGGTCTAAGATCCTTCCAGGCGCTTCAATGAGGAAATTT 1037
Qy 281 ValIleIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1038 GTGAAAAAATATGACTCATTGATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1097
Qy 301 ProCysGln 303
Db 1098 CCTGTGCAG 1106

RESULT 5
```

Db	763	TTTGAAGGTAGTAATATTACGACTGAAAAATGTGAAGCTTCAGCTGGGGTATTATTTCTTTGG	822
OY	221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	823	GAAAGCATACCCTCGTGAAACCCCTTGTATGAGATTGGGCCACCTTCGAAATCATG	882
OY	241	TTPAlaValHisAsnGlyThiArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	883	TGGGCTGTTCAATAAGTAGTACTGCAGCCACCATGATATAAATAATTACTTAAGCCCATGGG	942
OY	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	943	AGCCGATGACCTCGTTGTGTGCTAAAGAATCCCTCCACGCCCTTCATGTGAGGAATTT	1002
OY	281	ValLysIleMetThrHisIleuMetArgTyrrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	1003	GTGAAAATATATACCTTCATTGATGCGGTACTTCCAGGAGCACAGATGAGCCATTACAGTAT	1062
OY	301	ProCysGln	303
Db	1063	CCTGTGCAG	1071
RESULT 6			
AB009356			
LOCUS	AB009356	2769 bp	mRNA linear PRI 04-MAR-1998
DEFINITION	Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.		
ACCESSION	AB009356.1		
VERSION	AB009356.1 GI:2924623		
KEYWORDS	TAK1a; TGF-beta activated kinase 1a.		
SOURCE	Homo sapiens lung cDNA to mRNA, clone _lib:lambda gt11		

ORGANISM	REFERENCE
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
1 (sites)	
Sakurai, H., Shigemori, N., Hasegawa, K. and Sugita, T.	
TF-beta-activated kinase 1 stimulates NF-kappa B activation by an	
NF-kappa B-inducing kinase-independent mechanism	
Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)	
98153801	
2 (bases 1 to 2769)	
Sakurai, H.	
Direct Submission	
Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research	
laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Koshima-3-chome,	
Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp	
Tel:++81 6 300 2571, Fax:++81 6 300 2593)	

```

FEATURES
  source
    Location/Qualifiers
      1..2769
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="PBSTAXA1a"
        /issue_type="lung"
        /clone_idb="Lambda ST11"
        163..1902
          CDS

```

BASE COUNT	811 a	565 c	640 g	753 t
------------	-------	-------	-------	-------

ORIGIN	
Alignment Scores:	
Pred. No.:	4,48e-149
Score:	165.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
Length:	2769
Matches:	303
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-830-144-2_COPY_1_303 (1-303) x AB009356 (1-2769)

[illegible]


```

/db.xref="taxon:9606"
/cdname="pBstAK13"
/tissue_type="lung"
/cdname_1b="Lambda gt11"
163..1983

```

```

/product="TGf-beta activated kinase 1b"
/protein_id="BAA25026.1"
/db_xref="gi:2924626"
/translation="MSTSAASSSSSSAGMIAPSVLNFETIDKEIVVEYVGRG
GAFGVCAKMRKADVALKQIESSEKPAFLVEIRQUSRVNHPINVKLYGCLNPVCL
VMEVAEGSSLVKIDCFGAPDLYPTAAAHMSKCLQSOVALAHSMQPAFLHRDKKPP
NLLVAGGTVLTKIDCFGAPDLYPTAAAHMSKCLQSOVALAHSMQPAFLHRDKKPP
LMVEYITRRKPEDEIGGPAFRIMAAVHNTPRLPIKLNPKPIESLMTGCSMDQPPQSR
MEELVKITMLHMRFGPADEPLQYPCQYDGSQNSAISTSPMDIAWTSNDSNTNP
MEQVATNDTQIKRLSKLKNQAKQOQSSGSLISGASGSSVEELPSPSEKRSAM
SEIARIIAATTAAYGSKPKRKHGKTKASFGNIIAPVEIYISGNQPPRRSIQDLYVGT
GOVSRSSSPSVKMTTSGPTSEKPTSHPTPPDSTDITNSDMSIPMAVITLHDQLQ
PLACPNSKESMAVPEQHKQAQETMKQVETIALLQKQELVALLDQDEKQDQNTSR
LVQEHKLLDNKSLSTYYQOQCKELVIRSQOQKQDQTS"

```

Alignment Scores:	
Pred. No.:	4.64e-149
Score:	1615.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	100.00%
DB:	9
	Gaps: 0
	Length: 2850
	Matches: 303
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) X AB009357 (1-2850)

QY	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGluMetIleGlu	20
Db	163	AlrGtTAAGGCTTCGCCGCTCTCTCTCTCTGCTTCGGCGGGAGATGATTCAA	222
QY	21	AlaProSerGlnValIleuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu	40
Db	223	GCCCTTCACAGGCTCTCAACTTTGAAGAAGATCGATCAAGAGATCGAGTGGAAAG	282
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrPheAlaLysAspVal	60
Db	283	GTTGTGGAAAGAGGAGCTTTGGAGTTGTTGGCAAGCTTAAGTGGAGAGCAAAAGATGTT	342
QY	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	343	GCTTTTAAACAAATRGAAGTGAATCTGAAGAAAGGCTTATTTGTAAGCTTCGGAG	402
QY	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100
Db	403	TTATCCCGGTGGAAACCATCTTAATTTGTAAAGCTTTATGGAGCTGTTGAATCCAGTG	462
QY	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	463	TGCTTTGGAATGAATATGCTGAAGGGGCTTTATTAATATGCTGCATGAGTGCTGAA	522
QY	121	ProLeuProTyrTyrThrAlaAlaHisIleMetSerProCysLeuGlnCysSerGlnGly	140
Db	523	CCATTCGCCATTTTACTGCTGCCGCCAACCAGATGAGTGTGTTCACAGTGTCCCAAGGA	582
QY	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160
Db	583	GTGGCTTATCTTCACAGATGCCAACCCAAAGCGCTAATTCACAGGAGCTGAACACCA	642
QY	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	643	AACTTACTGCTGGTTGGCAGGGGAGACAGTTCTAAAAATTTGTGATTTGGTACAGCTGT	702
QY	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerIleAlaIleProGluVal	200
Db	703	GACATTGAGACACATGACCAATPACAGGGAGAGTGCTGTTGATGAGCACCTGGAAGTT	762

Qy	201	PheGluGlySerAspTyrSerGluLysGlyCysAspAlaPheSerTrpGlyIleIleLeuTrp	220
Db	763	TTTGAAAGTAACTAAATTACAGTGAAAAAAGTACACGCTTCACACTGGGATTAATTCCTTTGG	822
Qy	221	GIUValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaIlePheArgIleVal	240
Db	823	GAAAGTAAGAACGGCGCGGAAACCTTTGATAGATTGGTGCCAGCGCTTTCGGAATCATG	882
Qy	241	TryAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	883	TGGCGTGTTCAATTAAAGTACTGACCAACCACTGATTAATAAAATTTACTTAAGCCCATTTGAG	942
Qy	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	943	AGCCGATGACTCGTGTGGTGTCTTAAGATCTCTTCCAGCGCCCTTCAATGAGAGAAATT	1002
Qy	281	ValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	1003	GTAAGAAATAATGACTCACTTGATGCGGATCTTTCCAGAGACAGATGAGCCATTACAGTAT	1062
Qy	301	ProCysGln	303
Db	1063	CTTGTCAG	1071

RESULT 9	E38398	2866 bp	DNA	linear	PAT 31-JAN-2002
LOCUS	E38398				
DEFINITION	NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same.				
ACCESSION	E38398				
VERSION	E38398.1	GI:16626978			
KEYWORDS	JP 2000197500-A/4.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (Pases 1 to 2866)				
AUTHORS	Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.				
TITLE	NF-kappa B activation inhibitory drug targeting TAK1 and method for identifying the same				
JOURNAL	Patent: JP 2000197500-A 4 18-JUL-2000;				
COMMENT	TRANAB SEIVAKU CO LTD				
	OS Unidentified				
	PN JP 2000197500-A/4				
	PD 18-JUL-2000				
	PF 04-FEB-1999 JP 1999026803				
	PR				
PI	NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI HASEGAWA				
PC	C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC C12Q1/02.				
PC	G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09, PC C12R1.91).				
PC	C12N5/00,C12N15/00,(C12N15/00,C12R1.91)				
CC	Strandedness: Double;				
CC	Topology: Linear;				
PH	Key				
FT	source				
FT	1..2866				
	Location/Qualifiers				
	1..2866				
	Location/Qualifiers				
	1..2866				
	/organism='unidentified'				
	/db_xref="taxon:32644"				
BASE COUNT	848 a 588 c 656 g 774 t				
ORIGIN					
FEATURES					
source					
Alignment Scores:					
Pred. No.:	4.67e-149	Length:	2866		
Score:	1615.00	Matches:	303		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	100.00%	Indels:	0		
GB:	6	Gaps:	0		

US-09-830-144-2_COPY_1_303 (1-303) x E38398 (1-2866)		TAK1; TGF-beta-activated kinase; protein kinase.	
Qy	1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20	Mus musculus	
Db	163 ATGTCATCAGCGCTCGCGCGCTCCTCCTCCTCCTCGCTTCGCGCGGTGAGATCGAA 222	Mus musculus cDNA to mRNA.	
Qy	21 AlaProSerGlnValLeuAsnPheGluGluIleAspTyrIysGluIleValGluGlu 40	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	223 GCCCTTCCAGGTCCTCACTTTGAAGAGATCGACTACAGAGATCGAGGTGGAAG 282	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	283 GTTCTTGGAAGAGAGCGCTTTGAGTGTGTTGCAAGCTTAAGTGGAGACAAAGATGT 342	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	343 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAACGCTTATTGTAGAGCTTCGGCAG 402	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyValaCysLeuAsnProVal 100	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	403 TTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGGAGCTGCTGTAATCAGTG 462	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	463 TGTCTTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTGTCATGCTGTCAA 522	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	523 CCATTGCCATATTATCTGCTGCCAGCAATGAGTGTGTTTACAGTGTTCCTCAAGGA 582	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	583 GTGGCTTATCTTACAGCATGCAACCCAAAGCGCTAATTCACAGGACCTGAACACCA 642	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	643 AACTTACTGCTGGTTGCAGGGGGACAGTTCTAAATAATTGTGATTTGGTACAGCCTGT 702	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	703 GACATTCAGACACATATGACCAATAACAAGGGAGTGCTGCTTGGATGCGACCTGAAGTT 762	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeuLeuTrp 220	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	763 TTTGAAGGTAGTATTACAGTGAATAATGTGACGTCTTCAGCTGGGGTATTATCTTTGG 822	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	823 GAAGTGATACGGCTCGAAACCCCTTTGATGAGATTGGTGGCCAGCTTTCGAATCATG 882	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	883 TGGGCTGTTTATAATGTTACTGCACACCACTGATAAAAAATTTACCTAAGCCCATTCAG 942	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluIle 280	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	943 AGCTGTATGACTCGTTGTTGTTCTAAGATTCCTTCCACGCCCTTCAATGAGGAAATT 1002	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyValaAspGluProLeuGlnTyr 300	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	1003 GTGAAATAATGACTCACTTGATGCGGTACTTTCAGGAGCAGATGAGCCATTACAGTAT 1062	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Qy	301 ProCysGln 303	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
Db	1063 CCTGTCTAG 1071	Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;	
RESULT 10		MUSTAK1	
LOCUS		MUSTAK1	
DEFINITION		2443 bp mRNA linear ROD 04-FEB-1999	
ACCESSION		D76446	
VERSION		D76446.1 GI:1167505	


```

Db 448 TGTCTTGTGATGGAATATATGACAGGGGGCTCATTTGTATATATGCTGATGCTGCTGAA 507
Qy 121 ProLeuProTyrThrAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
Db 508 CATTCGCTTACTACTCTGCTCTATGCGATGAGCTGGTGTTCAGTGTTCACAAAGA 567
Qy 141 ValAlaTyrLeuHisSerMetGlnProTyrAlaLeuLeuHisArgAspLeuLysPro 160
Db 568 GTGGCTTACTGTCACAGCATGCGCCAAAGCGCTGATTCACAGGACCTCAAGCCTCCA 627
Qy 161 AnLeuLeuLeuValAlaGlyThrValLeuLysLeuCysAspPheGlyThrAlaCys 180
Db 628 AACTTCTCTCTGTTGTCAGAGGGACAGTTCTAAATAATCTGCGATTTGGTACACCTTGT 687
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluVal 200
Db 688 GACATCCAAACACACATGACCAATAATAAAGGAGTGCTGTGGATGGCGCTGAAGTG 747
Qy 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyrGlyIleLeuTyr 220
Db 748 TTTGAAGGTAGCAATTACAGTGAAGAGTGTGATGCTTTCAGCTGGGGTATTATCTCTCG 807
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 808 GAAGTGATAACAGCGCGGAAACCTTCGATGAGATCGGTGGCCAGCTTTTCAGAAATCATG 867
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuLeuLysAsnLeuProLysProIleGlu 260
Db 868 TGGGCTGTTCATATGCGACTCGACCACCATGATCAAAATTTTACCTAAGCCCATTTGAG 927
Qy 261 SerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 928 AGCTTGATGACAGCGTGTGGTCTAAGGACCATCTCAGCGCCCTTCAATGAGGAAATT 987
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 988 GTGAAATAATGACTCACTTGATGCGGTACTTCCAGGAGCGGATGAGCCATTACAGTAT 1047
Qy 301 ProCysGln 303
Db 1048 CCTTGTGAC 1056

RESULT 12
XLU92030 2812 bp mRNA linear VRT 18-APR-1998
LOCUS Xenopus laevis TGF-beta-activated kinase TAK1 mRNA, complete cds.
DEFINITION U92030
ACCESSION U92030
VERSION U92030.1 GI:3057035
KEYWORDS Xenopus laevis.
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis.
REFERENCE 1. 2812
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
Matsumoto,K., Nishida,E. and Ueno,N.
Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
EMBO J. 17 (4), 1019-1028 (1998)
MEDLINE 98130593
PUBMED 9463380
REFERENCE 2 (bases 1 to 2812)
AUTHORS Shibuya,H.
Direct Submission
TITLE Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES
source 1..2812
/organism="Xenopus laevis"
/db_xref="taxon:8355"
341..2191
/notes="TGF-beta-activated kinase"
/codon_start=1
CDS

```

```

/product="TAK1"
/protein_id="AAC14008.1"
/db_xref="GI:3057036"
/tranlation="MSATSAEMITPPVLNPEIDYKELEVEEVVGRGTFGVVCRAKW
RKDVAKIIESESEKAFIVELRQLSRVHNHIVKLYGACLNPLVCLWYEAEGSLY
NLVHGAEPYPTAAAHMSWLCQACQAYLHSMKPKALIHRLDLPVLLVAGGTVL
KICDGTACDIQTHMTNNKSPAAIIEGSLNTRCWSKDFVSGWIIWEVITRRKPF
DEIGFAPRIMWAVHNGTRPPIIKNLKPKIESLMTRCWSKDFVSGWIIWEVITRRKPF
MOVFPGADVLOPCQYDEGOSGATSTGSCIDITSTNTSKSDIHIEPDPQASAS
NNTIKIESKLAHLKNOAKOTGESGRSLSPSRGSSVESLSEIRGRPSTUGTSEGG
RMSADSELEARISASTAVYKPKGRHKRTAFSGNILDVPEIITAGNQOQERSVDL
TIVTESOESNRSSRSPSRVMTITSGTPDKPRGLWPAPDESDDTNGSDNIPM
AYLTDLHQLPLAPCNSKMAVFEQCMQAYKMQVTEIALILRQKQELIAELDQ
DEKQOQTSRLVQEHKKLLDENKSLSTYYQQCKQLEIVRSQQKQOQTS"
BASE COUNT 808 a 612 c 669 g 723 t
ORIGIN
Alignment Scores:
Pred. No.: 8,17e-137 Length: 2812
Score: 1490.50 Matches: 277
Percent Similarity: 96.92% Conservative: 6
Best Local Similarity: 94.86% Mismatches: 8
Query Match: 92.29% Indels: 1
DB: 5 Gaps: 1
US-09-830-144-2_COPY_1_303 (1-303) x XLU92030 (1-2812)
Qy 12 SerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGluGluIle 31
Db 344 TCTCCCACTCTGCTGAAATGATTGAGACTTCCA---CCGGTGTCTGAACCTTTGAGGAGATC 400
Qy 32 AspTyrLysGluIleGluValGluValGlyArgGlyAlaPheGlyValValCys 51
Db 401 GACTACAGGAGATTGAGGTGGAAGAGGTAGTTGGAAGAGAACCTTTGCTGTCTGC 460
Qy 52 LysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSerGluArg 71
Db 461 AAAGCTAAATGGCGAGGAAAGATGTGGCAATCAACAAATTTGAAAGCGAATCTGAAAGG 520
Qy 72 LysAlaPheIleValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLys 91
Db 521 AAAGCCTTTATTGTGTGAGCTGCGACACATTTGCTGCTGTAAACCATCCCAATATTGTCAAG 580
Qy 92 LeuTyrGlyAlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySer 111
Db 581 TTGTATGAGCGCTCCCTAAATCCGGTATGTTTAGTGTAGAGTATGCGAAGGAGGTTC 640
Qy 112 LeuTyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMet 131
Db 641 TTGTACAATGTTTTCATGGAGCTGAACCGCTGCTTACTATACAGCTGCTCATGCAATG 700
Qy 132 SerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAla 151
Db 701 AGTTGGTGTTTTACAATGTGCCAAGGAGTTGCATATTTTACATAGCATGAAGCCGAGGCT 760
Qy 152 LeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeu 171
Db 761 CTATTTACAGGAGCTCAACACCAACCACTTGTCTGTGTAGCTGGAGCACCCTTCTT 820
Qy 172 LysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGly 191
Db 821 AAGATTGTGACTTTTGTGTACAGCTGTGATATTACAGACTCACATGACTAATAACAAAGGA 880
Qy 192 SerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAsp 211
Db 881 AGTCAGCATGATGGGTGCCAGAGTTTTCAGAGGTAGCAATTACACCGAGAAATGTGAC 940
Qy 212 ValPheSerTrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGlu 231
Db 941 GTATTTAGTTGGGAAATAATTTCTTTGGAGAGTAAATAACCCGAGAAACCTTTCATGAA 1000
Qy 232 IleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu 251

```

Df 1001 ATTGGTGTCCAGCGTTCCTCATCATGTGGCGCTTGCACAAATGACTGCTCCACCATT 1060

Qy 252 11elyAsnleupProlyspProilegiUsrLeMeThArGysTyrSerlyssapPro 271
|||||
Df 1061 ATTAATAATTTGGCTTAAGCTATTGAAAGCTTAATGACTCGCTGGTGCACAAATGCC 1120

Qy 272 SerGlnArProSerMetGluGluileValyslleMerThHisleuMeKArgYrPhe 291
|||||
Df 1121 CCACAAGACCTTCATATGAGGAGATGTCTCAGATATATACATCATTAATGACGATATTTT 1180

Qy 292 ProGlyAlaAspGluProLeuGlnTyrProCysGln 303
|||||
Df 1181 CCTGGAGCGGATGTTCTTCACTATCTCTTGTCTG 1216

RESULT 13
AC114407 135147 bp DNA linear HTG 26-JUN-2002
AC114407
LOCUS
DEFINITION Mus musculus clone RP23-51G1, WORKING DRAFT SEQUENCE, 5 ordered pieces.
AC114407 GI:21592111
VERSION AC114407.3 GI:21592111
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULFILLTOP.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 135147)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Mus musculus, clone RP23-51G1
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 135147)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baetsen,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Charato,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Lander,E., Lehoczeky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Margus,N., Matthews,C., McCarthy,M., McKernan,K., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomas,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viet,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 135147)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Baetsen,V., Bloom,T., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K., Lamazares,R., Lander,E., Lehoczeky,J., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Margus,N., Matthews,C., McCarthy,M., McKernan,K., Meldrum,J., Menes,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomas,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viet,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE
JOURNAL
COMMENT
Direct Submission
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 26, 2002 this sequence version RepeatMasker: gi:21536031.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L23267
Center clone name: 51 G.1
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 13366 bases at least Q40
Consensus quality: 134318 bases at least Q30
Consensus quality: 134477 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 134747; sum-of-contigs
Quality coverage: 11.0 in Q20 bases; agarose-fp
Quality coverage: 12.3 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 12654: contig of 12654 bp in length
* 12655 12754: gap of 100 bp
* 12755 13467: contig of 733 bp in length
* 13468 13587: gap of 100 bp
* 13588 14337: contig of 750 bp in length
* 14338 14437: gap of 100 bp
* 14438 63285: contig of 48848 bp in length
* 63286 63385: gap of 100 bp
* 63386 135147: contig of 71762 bp in length.
Location/Qualifiers
1. 135147
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-51G1"
/clone_11b="RP23-51G1 Female Mouse BAC"
1. 12654
/note="assembly_fragment
clone end:SP6
vector_side:left"
misc_feature 12755..13467
/note="assembly_fragment"
misc_feature 13588..14337
/note="assembly_fragment"
misc_feature 14438..63285
/note="assembly_fragment"
misc_feature 63386..135147
/note="assembly_fragment
vector end:17
vector_side:right"
BASE COUNT 43081 a 25796 c 26087 g 39780 t 403 others
ORIGIN

Db 1033 GACAAAGCTGTTGCCGTCACAGAGATTCTTCGCCAGCGCCGAGACAGACATCGAGAAG 1092
Qy 77 GtUleuArGlnleuSerArgValaIasnHlsProAemIlleValleuLeuTyrlly----- 94
Db 1093 GAGGTAAACAGATTGTTCGGCCGCGGAGACCCGAAACATATGCTCTGCACGGGATATCC 1152
Qy 95 AlAcyleuAAsnProvalCysleuValMetGluTyralAGluGlyGlySerleuTyraSn 114
Db 1153 TCGTACCAGACAGCCCACTTACCTGATTAATGAGATTGCCGAAAGGTGATGCTGCACAAAC 1212
Qy 115 ValLeuHlsGlyAlAGluProleuProTyTyTyThralAlaHlsAlaMetSerTyPcys 134
Db 1213 TTCCTTCAAGCGC--AAGGTAAAGCCGGCATATTCTCTGCCCAAGCCCATGAGCTGGGGC 1269
Qy 135 LeuGlnCysSerGlnGlyValAlaTyLeuHlsSerMetGlnProlyAlaIleuIleHls 154
Db 1270 CGCAATGTCAGAGGCTGTGGCATTTTGACATGTCATGACCCCAAAACCATTAATACAT 1329
Qy 155 ATGAspleuLysProProAAsnLeuLeuValAlaGlyGlyTyThralValleuLysIleCys 174
Db 1330 CGGAGAGTGAAGCCGCTGAACCTGCTCTTGACCAACAGAGAGCCGCAATTCGAAGATATGC 1389
Qy 175 AspPheGlyThzAlaCysAspIleGlnThzHlsMetThzAsnAsnLysGlySerAlaAla 194
Db 1390 GACTTCGGCAGCGTGGCGGCAAGTCGACCATGATGACCAACATGCGCGTAGTGCCT 1449
Qy 195 TrpMetAlaProGluValAPheGluGlySerAsnTyTySerGluLysCysAspValPheSer 214
Db 1450 TGGATGGCCCGCCGAGGCTCTTCGAAAGGCTCCAAGTATACGAGAAAGTGTCACATTTTCAGC 1509
Qy 215 TrpGlyAlaIleleuTyrgluValAlaIleThzArgArgLysProPheAspGluIleGlyGly 234
Db 1510 TGGCCATTGTTTATAGGAGGTTCTGTCCAGAGACAGACCCCTTTAAAGCATTCGACAAAT 1569
Qy 235 ProAlaPheArgIleMetTyAlaValHlsAsnGlyThzArgProProleuIleLysAsn 254
Db 1570 ---GCCTACACCATCCAGTGAAGGAACTCAAGAGGTGAACGCCCGCGCTGCACCACT 1626
Qy 255 LeuProLysProIleGluSerleuMetThzArgCysTrpSerLysAspProSerGlnArg 274
Db 1627 TGGCCCAAGCGCATCGAGACCTGATGACCGCTGCGTGAAGAAACGGTGGCCAGAGATGC 1686
Qy 275 ProSerMetGluGluIleValLysIleMetThzHlsMetArgTyThzPheProGluAla 294
Db 1687 CGCTCATGAGTACATAGTGGCGCTTATGACAGAGATGTCAGAGACTATACGGGGGCG 1746
Qy 295 AspGluProLeuGlnTyTyr 300
Db 1747 GACAAAGCCCTGGAAATAC 1764

RESULT 15
AY051953 3386 bp mRNA linear INV 27-AUG-2001
LOCUS Drosophila melanogaster LD42274 full length cDNA.
DEFINITION AY051953
ACCESSION AY051953.1 GI:15292216
VERSION FL1_CDNA.
KEYWORDS Drosophila melanogaster.
SOURCE Drosophila melanogaster.
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 3386)
Stapleton, M., Briskestein, P., Hong, L., Aghayani, A., Carlson, J.,
AUTHORS Chang, M., Chavez, C., Dorsett, V., Farfan, D., Frise, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,
Nunoo, J., Pacleb, J., Paragas, V., Park, S., Phouanavong, S., Wan, K.,
Yu, C., Lewis, S. E., Rubin, G. M. and Ceiniker, S.
TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
COMMENT Sequence submitted by:

Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES
source
Location/Qualifiers
1..3386
/organism="Drosophila melanogaster"
/strain="Y; cn bw sp"
/db_xref="taxon:7227"
/map="19E1-19E1"
1..3386
/gene="CG18492"
/note="alignment with genomic scaffold AB003571"
/db_xref="FLYBASE:FBgn0042230"
935..2971
/gene="CG18492"
/note="Longest ORF"
/codon_start=1
/product="LD42274p"
/protein_id="AAK93377.1"
/db_xref="GI:15292217"
/translation="MATASLDALQAAVDFSEITLREKVGHSYGVNCAKVARDLIVA
VKEFASAEKDIKEVKQLSRVKNITIALGHSYQOATYLIWEPFGSGLNHPI
GKKRPAYSIALAHMSMAROCBEGLATYIHANTPRPLIHRDQKPLNLLTKGRNLKIDF
GVVADKSTMTNNRSGSAAMWAVEFEGSKYTEKDI FSAIVLMEVLSRKQPFKIDN
AVTIQWKLYKGRSPPLITCPRKIEDLMTACMKTYBEDRPSMQYIVGVNHEIVKYTG
ADKALRYTFVNOQIVTKESDGVAAQOPDLSQEGELSPSSQIDLPPTAAANVAIA
ISKTTSSMTENTSSTSDIPRNSGGLDNMLPFVMVNRMAIPIEESNIESRNSFN
LNSAATQRLRTIRNGTINAGACMKMEQITLDEVANGQVDRDELSSESSSTAKSGRE
RLTVITKRVMMTTDISNNNGIINHNSNLSHANGWQVDRDELSSESSSTAKSGRE
LVVDPRDENDGTQSLAEIIDLPELQPEPIPDNAEQIYRDRHMAKEYLSDVTNL
VYADQFKDLIVQMDRTEREQKQELRLKKIKDEGLQSLYNNLQOQVYASROLAAGHHPO
PHEHPHNPOLQHPHSPMHFLQDRGCGLLPQSVCGSSVSEGVVPIPHHNA"

BASE COUNT 956 a 898 c 857 g 675 t
ORIGIN
Alignment Scores:
Pred. No.: 1,56e-70 Length: 3386
Score: 818.00 Matches: 155
Percent Similarity: 69.93% Conservative: 45
Best Local Similarity: 54.20% Mismatches: 82
Query Match: 50.65% Indels: 4
DB: 3 Gaps: 3
US-09-830-144-2_COPY_1_303 (1-303) x AY051953 (1-3386)

Qy 17 GtUleuTlleguAlaProSerGlnValleuAAsnPheduGluIleAspTyrllyGluIle 36
Db 932 GAGATGGCCACAGATGCTGGAGCGACGCGACGACCTTATGACATTCGATGAGATA 991
Qy 37 GtUleuAlAGluValAlaGlyArgGlyAlaPheGlyValValaCysLysValaLysTrpArg 56
Db 992 ACACCTAAGAGAAAGTCGCGCCATGGGTCTTACGAGAGTGTTCGAAGCCGTTTGGCC 1051
Qy 57 AlAlaAspValAlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal 76
Db 1052 GACAAAGCTGTTGCCGTCACAGAGATTCTTCGACAGCGCCGACGAAAGCATCGAAG 1111

```
QY 77 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGly----- 94
Db 1112 GAGGTGAAGCAGTTGTCGCGGTGAAGCACCCGAAACATCATCGCTCTGCACGGGATATCC 1171
QY 95 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 114
Db 1172 TCGTACACAGCAGGCCACTTACTGATTAATGGAGTTCGCCGAAGGTGGATCGTGCAAC 1231
QY 115 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 134
Db 1232 TTCCTTCACGCGC--AAGGTGAAGCGGCATATTCTCTGCCCCACGCGCATGAGCTGGCG 1288
QY 135 LeuGlnCysSerGlnGlyValAlaTyrIleuHisSerMetGlnProLysAlaLeuIleHis 154
Db 1289 CGCCAATGTGCAGAGGGTCTGGCATATTTCATGCCATGACGCCAAAACCACTAATACAT 1348
QY 155 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 174
Db 1349 CGCAGCTGAAGCGCTGAACCTCTCTTGACCAACAGGACGCAATCTGAAGATATGC 1408
QY 175 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 1409 GACTTCGCGCACGGTGGCGGACCAAGTCGACCATGATGACCAACATCGCGCAGTGGCGT 1468
QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 1469 TGGATGGCGCCGAGGTCTTCGAAGGCTCCAAAGTATACGGAGAAGTGTGACATTTTAGC 1528
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
Db 1529 TGGCCCATTTGTTCTATGGAGGTTCTGTCCAGGAAGCAGCCCTTAAAGGCATCGACAAT 1588
QY 235 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1589 --GCCTACACCATCCAGTGGAGATCTACAAGGTGAACGCCCGCGCTGCTGACCACT 1645
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATACCGCCTGTCTGGAAAAACGGTGCCTCGAGGATCGC 1705
QY 275 ProSerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGlyAla 294
Db 1706 CCGTCGATGCAGTACATAGTGGCGGTATGCAGAGATCGTCAAGGACTATACGGGGCGC 1765
QY 295 AspGluProLeuGlnTyr 300
Db 1766 GACAAAGGCCCTGGAATAC 1783
```

Search completed: December 10, 2002, 03:06:37

Job time : 3127 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 9, 2002, 23:14:54 ; Search time 326 Seconds
(without alignments)
2093.115 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MTASASASSSSSSAGEMIE.....MTHLMRYFPFGADEPLQYPCQ 303

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame.p2n.model -DEV=xlp
-Q/cgn2_1/USFTO_spool/US09830144/runat_04122002_141510_2901/app_query.fasta_1.455
-DB=N_geneseq_101002 -QWTF=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USFR=US09830144 @CNG 1.1.79 runat_04122002_141510_2901 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1615	100.0	1704	20	AAK99698	Human TGF-beta act
2	1615	100.0	1788	20	AAK56285	Human TAK1-6xHis e
3	1615	100.0	1959	18	AAK85095	Human transforming
4	1615	100.0	2656	20	AAK56279	Human TAK1 encodin
5	1615	100.0	2656	21	AAK39105	Human TAK-1 nucleo
6	1615	100.0	2769	24	ABL88437	Pain regulated cDN
7	1615	100.0	2785	20	AAK99696	Human TGF-beta act
8	1615	100.0	2866	20	AAK99697	Human TGF-beta act
9	1609	99.6	2443	18	AAK85094	Mouse transforming
10	818	50.7	3367	23	ABL02489	Drosophila melanog
11	689.5	42.7	10997	23	ABL02488	Drosophila melanog
12	505	31.3	759	23	ABL08337	Drosophila melanog
13	482.5	29.9	3454	24	ABL70018	Pancreas cancer re
14	445	27.6	3141	22	AAK18824	Human kinase (PKIN
15	443	27.4	3538	24	AAK34309	Human PKIN-12 cDNA
16	442	27.4	2120	21	AAK299726	cDNA encoding huma
17	434.5	26.9	1224	21	AAK43254	Arabidopsis thalia
18	434	26.9	1063	22	AAH34976	Human colon cancer
19	434	26.9	1631	21	AAK39537	Arabidopsis thalia
20	434	26.9	1706	21	AAK75674	DNA encoding a hum
21	434	26.9	2191	22	AAH99263	Human protein enco
22	434	26.9	2194	21	AAK299734	Cardiovascular sys
23	434	26.9	2220	21	AAK293783	Human survival reg
24	434	26.9	2622	22	AAK75336	Human TGF-beta rec
25	434	26.9	3967	22	AAH73366	Human cervical can
26	433	26.8	1365	21	AAK299727	cDNA encoding huma
27	433	26.8	2272	21	AAK299736	Cardiovascular sys
28	433	26.8	2403	22	AAK44701	Novel protein kina
29	432.5	26.8	3066	24	ABQ86185	Novel human gene.
30	432	26.7	2157	22	AAH46913	cDNA encoding huma
31	432	26.7	3111	24	ABN86357	Novel human protei
32	432	26.7	3518	24	ABN86358	Novel human protei
33	427.5	26.5	1591	21	AAK40839	Arabidopsis thalia
34	427.5	26.5	3072	23	ABL04355	Drosophila melanog
35	419	25.9	3558	24	ABK83874	Human cDNA differe
36	419	25.9	3558	24	AAK36139	Human mitogen acti
37	415.5	25.7	3063	23	ABL08336	Drosophila thalia
38	411.5	25.5	2283	21	AAK48526	Arabidopsis thalia
39	410	25.4	2505	21	AAK47607	Human CARK (Cardia
40	410	25.4	2505	21	AAK47609	Rat CARK (Cardiac
41	410	25.4	2508	22	AAK44702	Novel protein kina
42	410	25.4	2788	23	AAK80720	DNA encoding novel
43	410	25.4	3025	21	AAK47606	Human CARK (Cardia
44	410	25.4	3026	21	AAK47608	Rat CARK (Cardiac
45	401	24.8	3365	24	ABK84203	Human cDNA differe

ALIGNMENTS

RESULT 1

AAK99698

ID AAK99698 standard; cDNA to mRNA; 1704 BP.

XX

AC AAK99698;

XX

DT 18-OCT-1999 (first entry)

XX

DE Human TGF-beta activated kinase (TAK) 1c encoding nucleotide sequence.

XX

KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW endotoxin shock; septicemia; human; hTAK1c; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

```
FT CDS 1..1704
FT /tag= a
FT /product= "hTAK1c"
PN WO9940202-A1.
PD 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-JP00422.
XX
XX 30-OCT-1998; 98JP-0309316.
XX 06-FEB-1998; 98JP-0026003.
XX
XX (TANA ) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
PI
XX
XX WPI; 1999-494298/41.
DR P-PSDB; AAY28998.
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases
XX
XX Examples; Page 43-46; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the nucleotide sequence of human TAK1c
CC (hTAK1c) protein.
XX
XX
SQ Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;

Alignment Scores:
Pred. No.: 2..19e-173 Length: 1704
Score: 1615..00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x AAX99698 (1-1704)
QY 1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGlu 20
DB 1 ATGTCTACAGCCTCTGCGCCCTCTCTCTCTCTCTGCGCCGCTAGATGATCGAA 60
QY 21 AlaProSerGlnValLeuAsnPhenGluGluIleAspTyrLysGluIleGluValGluGlu 40
DB 61 GCCCTTCCAGAGTCTCTCAACTTTGAAGATGATGATCAACAGAGATGAGTGAAGAG 120
QY 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrrArgAlaLysAspVal 60
DB 121 GTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
QY 61 AlaIleLysGluIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 181 GCTATTAAACAATTAAGAACTGATCTGAGAGAAAGCCTTTATTAGAGCTTGGCGAG 240
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
DB 241 TTATCCCGTGTGAACCATCTTAATATTGTTAAAGCTTTATAGAGCCTTGATGATCAG 300
QY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
DB 301 TGTCTTGTGATGAAATGCTGAAGGGGGCTTTTATATATGCTGCATGCTGCTGAA 360
QY 121 ProLeuProTyrTrrThrAlaAlaHisAlaMetSerTrrCysLeuGlnCysSerGlnGly 140
```

```
DB 361 CCATTGCCATATATATACTGCTGCCACAGCAATGAGTGTGTTTACAGTGTGCCAAGA 420
QY 141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
DB 421 GTGGCTTATCTTCACACAGACGCAACCCAAAGCCCTAATTACAGGAGACCGAAACACA 480
QY 161 AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
DB 481 AACTTACTGCTGTTGACAGGGGGGACAGTCTTAAAAATTTGTATTTTGTATACGCTGT 540
QY 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIleTrrMetAlaProGluVal 200
DB 541 GACATTCAGACACACATGACCAATTAACAAGGGAGTGTCTCTTGATGATGACCATGAA 600
QY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrGlyIleIleLeuTrr 220
DB 601 TTTGAAGTATGTAATTACAGTGAATAATGACGCTCTTCAGCTGGGGTATATTCTTTGG 660
QY 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
DB 661 GAAGTATTAACGCTCGGAAACCTTTGATGAGATTGTGGCCAGCTTTCCGAATCATG 720
QY 241 TrrAlaValHisAsnGlyTrrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
DB 721 TGGCTGTTCATATATGTTACTCGACCACTGATAAAAATTTACTTAAGCCCATTTGAG 780
QY 261 SerLeuMetThrArgCysTrrSerLysAspProSerGlnArgProSerMetGluGluIle 280
DB 781 AGCTGTATGACTCTTGTGTGTAAAGATCTTCCAGCGCCCTTCATATGAGAGAAATT 840
QY 281 ValLysIleMetThrHisLysMetArgTyrPheProGluValAlaAspGluProLeuGlnTyr 300
DB 841 GTGAATAATATGACTCTCACTTATGCGGTACTTTCCAGAGCAAGATGAGCCATTACAGTAT 900
QY 301 ProCysGln 303
DB 901 CTTGTCTAG 909

RESULT 2
AAX56285
ID AAX56285 standard; DNA; 1788 BP.
XX
XX AAX56285;
AC
XX 21-JUL-1999 (first entry)
DT
XX
XX Human TAK1-6xHis encoding DNA.
DE
XX
XX Human; TAB1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
XX Homo sapiens.
OS
XX synthetic.
XX
XX Key Location/Qualifiers
FT CDS /tag= a
FT
XX
XX WO9921010-A1.
XX
XX 29-APR-1999.
XX
XX 22-OCT-1998; 98WO-JP04796.
XX
XX 22-OCT-1997; 97JP-0290188.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Ohnomo T, Ono K, Tsuchiya M;
PI
XX
XX WPI; 1999-312645/26.
XX
XX P-PSDB; AAY09547.
XX
```



```

Db      363 GCTATTAAACAAATAGAGTGAATCTGAGAGAAAGCGTTATTGTGAGAGCTTCGGCAG 422
QY      81 LeuSerArgValAsnHisProAsnIleValIysLeuTYrGlyAlaCysLeuAsnProVal 100
Db      423 TTAATCCCGTGTAAACCATCTTAATTTGAAGCTTTATGGAGCCTGCTGTGATCAATG 482
QY      101 CysLeuValMetGluTYrAlaGluGlyGlySerLeuTYrAsnValIleHisGlyAlaGlu 120
Db      483 TGTCTTGATGGAATATCTGAAAGGGGCTCTTATATATATGTCCTGATGTCCTGAA 542
QY      121 ProLeuProTYrTYrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db      543 CCATGGCCCATATATATACGTGCTCCACGCAATGAGTGGTGTTCAGATGTTCCCAAGGA 602
QY      141 ValAlaTYrLeuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProPro 160
Db      603 GTGGCTTAATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGGAGCCTGAAACCA 662
QY      161 AsnLeuLeuLeuValAlaGlyGlyThrValIleuIysIleCysAspPheGlyThrAlaCys 180
Db      663 AACTTACTGCTGCTGGTCAGGGGGCAGATTCTAAAAATTTGTGATTTTGTACAGCCTGT 722
QY      181 AspIleGlnThrHisSerMetThrAsnLeuIysGlySerAlaAlaTrpMetAlaProGluVal 200
Db      723 GACATTCAACACACATGACCAATTAACAAGGGAGTGTGCTTGGATGGACCTGAAGTT 782
QY      201 PheGluGlySerAsnTYrSerGlnIysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db      783 TTTGAAGGTAGTATTAACGTAAAAATGTGACGTCTCAGCTGGGGATTTATCTTTGG 842
QY      221 GluValIleThrArgAlaGlySerProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db      843 GAAGTATTAACCGCTGGAAACCTTTGATGATGTGGTGGCCCAAGCTTTCCGAATCTATG 902
QY      241 TrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeuProIysProIleGlu 260
Db      903 TGGGCTGTTCAATATGTAAGTCTGACCAACCATGATAAAAAATTTACTTAAGCCCATTTG 962
QY      261 SerLeuMetThrArgCysTrpSerIysAspProSerGlnArgProSerMetGluGluIle 280
Db      963 AGCTTATGATGCTGTTGTTGTTAAGATCTTCCACAGCCCTTCATGAGGAAATTT 1022
QY      281 ValIysIleMetThrHisIleuMetCArgTYrPheProGlyAlaAspGluProLeuGlnTYr 300
Db      1023 GTGAAATATATGACTCATCTGATGCGGTACTTCCAGAGCAGATGAGCATTAAGAT 1082
QY      301 ProCysGln 303
Db      1083 CTTGTTCAG 1091

```

```

RESULT 6
ABL88437
ID      ABL88437 standard; cDNA; 2769 BP.
XX
AC      ABL88437;
XX
DT      16-MAY-2002 (first entry)
XX
DE      Pain regulated cDNA sequence 80.
XX
KW      Pain; analgesic; gene therapy; neurological disorder;
KW      neurodegenerative disease; gene; ss.
XX
OS      Homo sapiens.
XX
PN      WO200212338-A2.
XX
PD      14-FEB-2002.
XX
PF      03-AUG-2001; 2001WO-EP09011.
XX
PR      03-AUG-2000; 2000DE-1037759.
XX

```

```

PA      (CHEF ) GRUENTHAL GMBH.
XX
XX      Gillen C, Metzels I, Wendt S, Weihe E, Schaefer MK;
PI      WPI; 2002-257469/30.
XX
DR      P-PSDB; ABB85033.
XX
PT      Identifying pain-regulating compounds, useful for treating chronic pain
PT      and for diagnosis, by measuring binding of compounds to specific
PT      peptides and proteins
XX
XX      Claim 1; Fig 44; 213pp; German.
PS
CC      The invention relates to identifying pain-regulating substances (A)
CC      comprises (i) incubating a test substance with a cell (or preparation
CC      from it) that has synthesised a peptide or protein (B) and (ii) measuring
CC      either binding of the test substance to (B) or some functional parameter
CC      that is altered by this binding. The method is useful for identifying
CC      pain-regulating substances (A) with analgesic activity. (A) along with
CC      nucleic acid (ABL88411-ABL88441) that encode proteins (B).
CC      ABB85006-ABB85037) that interact with (A); (B); vectors containing the
CC      nucleic acid; antibodies against (B); cells that express (B) and agents
CC      that bind to (B), are all useful for treating pain, particularly chronic
CC      pain, including use in gene therapy. The same materials can also be used
CC      for diagnosis, e.g. of neurological and neurodegenerative diseases. The
CC      present sequence is that of a polynucleotide of the invention.
XX
SQ      Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;
XX
Alignment Scores:
Pred. No.: 4 5e-173 Length: 2769
Score: 1635.00 Matches: 303
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-2_COPY_1_303 (1-303) x ABL88437 (1-2769)
QY      1 MetSerThrAlaSerAlaAlaSerSerSerSerSerAlaGlyIleMetIleGlu 20
Db      163 ATGTCTACAGCTCTGCGCGCTCCTCCTCCTCGTTCGCGCGGTGAGATGCAAA 222
QY      21 AlaProSerGlnValIleuAsnPheGluGlnIleAspTYrIleGluIleGluValIleGlu 40
Db      223 GCCCCTTCCAGGCTCCCACTTGAAGATGACATCAAGAGATCGAGTGGAGAG 282
QY      41 ValValGlyArgGlyAlaPheGlyValIleCysIleAlaIleSTrpaAlaIysAspVal 60
Db      283 GTTGTGGAAGAGAGAGCCTTGGAGTGTGTTGCCAAAGCTAAGTGAAGCAAAAGATGTT 342
QY      61 AlaIleYsgIleGlnIleGlnSerGlnSerGlnArgIysAlaPheIleValGluLeuArgIle 80
Db      343 GCTATTAAACAAATAGAAAGTGAATCTGAGAGAAAGCGCTTTATTGAGAGCTTCGGCAG 402
QY      81 LeuSerArgValAsnHisProAsnIleValIysLeuTYrGlyAlaCysLeuAsnProVal 100
Db      403 TTAATCCCGTGTAAACCATCTTAATTTGAAGCTTTATGAGCCTGCTGAAATCCAGTG 462
QY      101 CysLeuValMetGluTYrAlaGluGlyGlySerLeuTYrAsnValIleHisGlyAlaGlu 120
Db      463 TGTCTTGATGGAATATCTGAAAGGGGCTCTTATATATATGTCCTGATGTCCTGAA 522
QY      121 ProLeuProTYrTYrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly 140
Db      523 CCATGGCCCATATATATACGTGCTCCACGCAATGAGTGGTGTTCAGATGTTCCCAAGGA 582
QY      141 ValAlaTYrLeuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuIysProPro 160
Db      583 GTGGCTTAATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGGAGCCTGAAACCA 642
QY      161 AsnLeuLeuLeuValAlaGlyGlyThrValIleuIysIleCysAspPheGlyThrAlaCys 180

```



```

XX Venter JC, Adams M, Li PMD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58385.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1949; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB57737-AB572072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2.74e-82 Length: 3367
Score: 818.00 Matches: 155
Percent Similarity: 69.93% Conservative: 45
Best Local Similarity: 54.20% Mismatches: 82
Query Match: 50.65% Indels: 4
DB: Gaps: 3
US-09-830-144-2_COPY_1_303 (1-303) x ABL02489 (1-3367)
OY 17 GJumctileglualaproserglnvalleuasnphleglugluileaspyrlysgluile 36
DB 932 GAGATGGCCACAGCATCGCTGGAGCCACTGCAGGAGCTTATGTGACTTACAGAGATA 991
OY 37 GIUValIGUluValValIGlYARglYAlapheglYValValCysLYsAlALysTrpArg 56
DB 992 ACACrPAAAGAGAAAGTGGCGCATGGCTCAAGGAGTGGCTGGCAAGCCGTTGGGCG 1051
OY 57 AlAlYsAspValAlAlAlLysglInleglUserglUsrGLuArgLYAlALpHeIleVal 76
DB 1052 GACAAAGCTGTTCCTCAAGAGATTCTTCGCGACGCCGAGCAGAAAGACATCGAGAA 1111
OY 77 GlUeuArGlnleUserArGValAsnHIsPrOAsnleValLYsLeuTrGly----- 94
DB 1112 GAGGTGAAGCAGTGTTCGCGCGTGAAGCACCCGAAACATCATCGCTTCGACGAGATATCC 1171
OY 95 AlAcYsLeuAsnProValCYsLeuValMetGLUTrAlagLUglYglYserleUTrAsn 114
DB 1172 TCGTACCAAGCCAGCCACCTACTGATATGAGATTGCGCGAAGGTGGATCGTCACAA 1231
OY 115 ValIeuHISglYAlagLUProleuProTrYrThrAlaAlAHISAlAMeSerTrpCys 134
DB 1232 TTCCTTCACGGC---AAAGTAAAGCCGCAATTCCTTCGCGCCAGCAGATGAGCTGGGCG 1288
OY 135 LeuGlnCYsSerGlnGlyValAlAlTYrIeuHISserMetGlnPRLYsAlALeuIleHIS 154
DB 1289 CGCCAAATGTGAGAGAGGTCTGCATATTTCATGCATGACCCAAAACCACTAATATCAT 1348
OY 155 ArgAspLeuLYsPrOProAsnleUleuValAlagLYglYThrValleUysIleCys 174
DB 1349 CGCAGCGTGAAGCCGCTGAACCTCTCTTGACCAACAGGAGCCCAATCGAAGATATGC 1408
OY 175 AspPheglYThrAlAcYsAspIleGlnThHISerThrAsnleUysglYserAlAla 194
DB 1409 GACTTCGGCAGCGTGGCGGACAACTGCATGATGACCAACAAATCGCGGAGGCGGCT 1468
OY 195 TrpMetAlaProGluValPheGluGlySerAsnTrYserGlnUysCYsAspValPheSer 214

```

```

DB 1469 TGAATGGCGCCCGAGGCTTCGAGAGCTCCAAAGTATACGAGAGATGTGACATTTTAC 1528
OY 215 TrpGlyIleIleleUeuTrpGluValIleThrARgArGLysPrOphesglutIleGly 234
DB 1529 TGGGCACTTGTCTTAAGGAGGTTCTGTCCAGGAGAGGCCCTTTAAAGCACTGACAA 1588
OY 235 ProAlapheArGlnleUeTrAlaValAlAsnGlyThrArpProleuIleUysAsn 254
DB 1589 ---GCCrPAAACACATCCAGTGGAGATCTACAAGGTGAAGCCCGCGCTGTCGACCACT 1645
OY 255 LeuProLYsPrOleglUserleUmetThrArgCYeTrpserLYsAspProserGlnArg 274
DB 1646 TGCCCCAAGGCGCATCGAGGACCTGATGACCGCTGTGMAACCGTGGCCGAGATCCG 1705
OY 275 ProserMetluGluIleValLYsIleMetThrIleUeuArGTrYrPheProGluYAla 294
DB 1706 CCGTCGATGCAATACATACATAGTGGCGCTTATGCACGAGATGCTCAAGACTATACGGGGCG 1765
OY 295 AspGluProleuGlnTrYr 300
DB 1766 GACAAAGCCCTGGAATAC 1783
RESULT 11
ID ABL02488 standard; cDNA; 10997 BP.
XX
XX ABL02488;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58385.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1946; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB57737-AB572072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;
XX

```


Best Local Similarity: 40.57% Mismatches: 71
Query Match: 31.27% Indels: 26
DB: 23 Gaps: 4

US-09-830-144-2_COPY_1_303 (1-303) x ABL08337 (1-759)

```
OY 29 GIUGLUIleAspTyrIleGluValIleGluValIleValIleArgIleValIlePheGly 48
DB 16 GAGGCTGTTCCCTTATGAGGATCCAGACAAAGACCTTATAGGCGCAGATTATGAGC 75
OY 49 ValIleCysIleValIleAspIleValIleValIleValIleValIleValIleValIle 68
DB 76 AGGTTTACAGGGCGGCTCGAGAAATCTGAGATCGCGCTGAGAGAGATTCGAGGGA 135
OY 69 SerGluArgIleValIlePheIleValIleGluLeuArgIleLeuSerArgIleValIleProIle 88
DB 136 TGTGAGAGTAAAAAGATTGAAAGAGATCTACAGCTACAAAGGCCAGCATGTCAAC 195
OY 89 IleValIleLeuTyrIleValIleValIleValIleValIleValIleValIleValIle 106
DB 196 ATAGTTGAGCTCTACCGCATCGACATCGAGCGACGAGGATGCGCCCTGCTGTGATGAAATTC 255
OY 107 AlaGluGlyIleSerIleTyrAsnValIleValIleValIleValIleValIleValIleValIle 126
DB 256 GTAGACGCTGATCTCTCTCCAGTTTCTGCAC---GGCAAAAGCAAGCCAGTTATTTCG 312
OY 127 AlaIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 146
DB 313 CATGCCACGCTTCAACATGCGCGCATCGATCGCTCAGGCGATGAGCTATCTGCATGCGC 372
OY 147 MetGluIleProIleValIleValIleValIleValIleValIleValIleValIleValIle 166
DB 373 ATGAGCGCCAAAGACGATTTATCGCGATTAAGCCACCTCAAAATACCTGATATGCGAG 432
OY 167 GlyIleIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 186
DB 433 AAGGAGCTCAACTGAGATTTGGCATTTGGCACTGTTTGGACTATCCCAATCGATA 492
OY 187 ThrAsnAspIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 199
DB 493 TCGTCATATGCGGACCGCATCGATACAAAGCGCGGAGTGAAGGAGTTGTTGATTTTC 552
OY 200 -----ValPheGluGlySer 204
DB 553 AATATCATCGATATATATATATACCAACCAACCGGCTTCAAAAGTTCTCAAGGAAT 612
OY 205 AsnTyrSerGluIleValIleValIleValIleValIleValIleValIleValIleValIleValIle 224
DB 613 AAACCCGATGAAAAGTGGATGTATATGTTGGCTATTACCTTTGGGAATATTTTCG 672
OY 225 ArgAlaGlyIleProIleAspGluIleGlyIleProIlePheArgIleMetTyrIleValIleHis 244
DB 673 CGCAAGGAGCCATTGACCAATAT---AATACGCTTTTGAACGTATACATGCTATTAAAT 729
OY 245 AsnGlyThrArg 248
DB 730 GAAGGCAAGAGA 741
```

RESULT 13
ABL70018
ID ABL70018 standard; DNA; 3454 BP.

AC ABL70018;

DT 15-MAY-2002 (first entry)

DE Pancreas cancer related gene sequence SEQ ID NO:8355.

KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW Stomach; lung; prostate; pancreas; carcinoma; antitumour; carcinos;
KW Cystostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.

OS Homo sapiens.
XX WO200194629-A2.

PD 13-DEC-2001.

PF 30-MAY-2001; 2001WO-US10838.

PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-233133P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236033P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
PR 28-SEP-2000; 2000US-236111P.
PR 28-SEP-2000; 2000US-236422P.
PR 29-SEP-2000; 2000US-236491P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.

(AVAL-) AVALON PHARM.

PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S,
PI Soppet DR, Weaver Z;

DR WPI; 2002-188264/24.

PT Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX
XX Claim 1; SEQ ID 8355; 44pp; English.

XX The present invention describes a method (M1) for screening for an
XX anti-neoplastic agent. The method involves exposing cells to a chemical
XX agent to be tested for anti-neoplastic activity, determining a change in
XX expression of at least one gene (I) of a signature gene set, where (I)
XX comprises a sequence (S) selected from 8447 sequences (given in ABL61664
XX to ABL70110), or is at least 95% identical to (S), where a change in
XX expression is indicative of anti-neoplastic activity. (I) has cystostatic
XX activity and can be used in gene therapy. M1 can be used for screening
XX an anti-neoplastic agent, and can be used for producing a product which

CC is the data collected with respect to the anti-neoplastic agent as a
 CC result of M1, and the data is sufficient to convey the chemical
 CC structure and/or properties of the agent. M1 can be used in the
 CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
 CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
 CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
 CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
 CC carcinoma, papillary carcinoma and Wilms' tumour.
 XX

SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;

Alignment Scores:

Pred. No.:	4.14e-44	Length:	3454
Score:	482.50	Matches:	107
Percent Similarity:	56.63%	Conservative:	51
Best Local Similarity:	38.35%	Mismatches:	98
Query Match:	29.88%	Indels:	23
DB:	24	Gaps:	7

US-09-830-144-2_COPY_1_303 (1-303) x ABL70018 (1-3454)

QY	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGlu	40
DB	535	GCACCCGGGCGCTCCAGCTGCCCGAGAGATCCCTTCACGAGCTGCAGCTAGAGGAG	594
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTyrArgAlaLysAspVal	60
DB	595	ATCATCGTGTGGGGGCTTTGGCAAGTCTATCGGCCCTGTGGCGTGGCGAGGAGTG	654
QY	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal-----	76
DB	655	GCAGTCAAGGCGCGCGCTGGACCTCGAAGAGGACCGCGCAGTCGACGCGGAGCAGGTG	714
QY	77	-----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly	94
DB	715	TGCCAGGAAGCCCGGCTCTTTGGAGCCCTGCAGCACCCCAACATAATTCGCCCTTAGGGGC	774
QY	95	AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu	112
DB	775	GCCTGCTCAACCCCGGCACACCTCTGCTAGTAGTGAGTATGCCCGGGTGTGCACTG	834
QY	113	TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer	132
DB	835	AGCAGGGTGTGGCAGGTGCGGGGTGCCACCTCACGTG-----CTGGTCAAC	882
QY	133	TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu	152
DB	883	TGGGCTGTGAGGTGGCGGGGCGATGAATACCTACACAATGATGCCCTGTGCCCATC	942
QY	153	IleHisArgAspLeuLysProProAsnLeuLeuVal-----	165
DB	943	ATCCACCGGACCTCAAGTCCATCACTCTGATCTGGAGGCCATCGAGAACCAAC	1002
QY	166	AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp-----IleGlnThr	184
DB	1003	CTCGCAGACACGGTGTCTCAAGATCACGGACTTTCGGCTCGCCCGCGAGTGGCACAGACC	1062
QY	185	HisMetThrAsnAnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer	204
DB	1063	ACCAAGATGAGCGCTGGGGGACCTACGCTGGATGGCGCGGAGGTTATCCGCTCTCTCC	1122
QY	205	AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr	224
DB	1123	CTCTTCTCCAAACAGCAGTATGTCTGGAGCTTCGGGGTGTCTGTGGAGCTGCTGACG	1182
QY	225	ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal---	243
DB	1183	GGGGAGGTCCCTTACCGTGAGATC-----GACGCTTGGCGGTATGGCGTGGCT	1236
QY	244	HisAnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet	263
DB	1237	ATGAATAAGTGCAGCTGCCCATTCCTCCACGTGCGCCCGAGGCCCTTTGCGCCGCTCTG	1296

QY	264	ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys	282
DB	1297	GAGGAATGCTGGACCCAGACCCCGCGCGCCAGATTTCCGTTAGCATCTTGAAG	1353

RESULT 14

AAAD18824	standard; cDNA; 3141 BP.
ID	AAAD18824
XX	AAAD18824;

DT	18-DEC-2001	(first entry)
----	-------------	---------------

DE	Human kinase (PKIN)-9 cDNA.
----	-----------------------------

KW	Human kinase; PKIN; gene therapy; adenocarcinoma; immune disorder; gout;
KW	cancer; allergy; sarcoma; leukaemia; acquired immune deficiency syndrome;
KW	AIDS; Addison's disease; microbial infection; inflammation; osteoporosis;
KW	atherosclerosis; cardiovascular disease; myocardial infarction; aneemia;
KW	myasthenia gravis; cirrhosis; cataract; growth and development disorder;
KW	seizure disorder; pulmonary embolism; Gaucher's disease; lipid disorder;
KW	lipid storage disease; Pick's disease; Tay-Sachs disease; renal disease;
KW	obesity; restorative therapy; immunomodulatory; vaccine; cardiovascular;
KW	antimicrobial; cytostatic; antiinflammatory; asthma; ss.

OS	Homo sapiens.
----	---------------

XX	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	1..3141
----	-----	---------

FT	FT	/*tag= a
FT	FT	/product= "Human PKIN-9 protein"

PN	WO200181555-A2.
----	-----------------

PD	01-NOV-2001.
----	--------------

XX	20-APR-2001; 2001WO-US12992.
----	------------------------------

XX	20-APR-2000; 2000US-199021P.
----	------------------------------

PR	28-APR-2000; 2000US-200226P.
----	------------------------------

PR	05-MAY-2000; 2000US-202339P.
----	------------------------------

PR	11-MAY-2000; 2000US-203505P.
----	------------------------------

PR	18-MAY-2000; 2000US-205564P.
----	------------------------------

PR	26-MAY-2000; 2000US-207739P.
----	------------------------------

PR	01-JUN-2000; 2000US-208795P.
----	------------------------------

XX	(INCY-) INCYTE GENOMICS INC.
----	------------------------------

XX	Yue H, Gandhi AR, Tribouley CM, Kearney L, Griffin JA, Nguyen DB;
----	---

PI	Bandman O, Lu DAM, Lal P, Burford N, Khan FA, Wallia NK, Yao MG;
----	--

PI	Patterson C, Burrill JD, Marcus GA, Zingler KA, Recipon SA, Lu Y;
----	---

PI	Policky JL, Thornton M, Tang YT, Hafalia A, Elliott VS, Baughn MR;
----	--

PI	Walsh RT, Ramkumar J, Borowsky ML, Au-young J, Hillman JL;
----	--

PI	Gururajan R;
----	--------------

XX	WPI; 2001-611740/70.
----	----------------------

DR	P-PSDB; AAE11775.
----	-------------------

XX	Human kinases and nucleic acids, useful for preventing diagnosing and
----	---

PT	treating cancers, inflammation and immune disorders -
----	---

XX	Claim 5; Page 157-158; 166pp; English.
----	--

XX	The present invention relates to human kinases (PKIN) and the nucleic
CC	acids encoding them. PKIN is used as vaccine and in gene therapy. PKIN is
CC	used in the prevention, diagnosis and treatment of diseases cancers,
CC	adenocarcinoma, leukaemia, sarcoma, immune disorder, Addison's disease,
CC	acquired immune deficiency syndrome (AIDS), aneemia, asthma, allergies,
CC	gout, microbial infections, cardiovascular disease and/or inflammation,
CC	myasthenia gravis, atherosclerosis, cirrhosis, osteoporosis, myocardial
CC	infarction, cataract, growth and development disorder, seizure disorder,
CC	pulmonary embolism, Gaucher's disease, lipid disorder, lipid storage
CC	disease, Pick's disease, Tay-Sachs disease, renal disease and obesity.
CC	PKIN may be used to treat disorders associated with decreased PKIN

expression by rectifying mutations or deletions in a patient's genome that affect the activity of PKIN by expressing inactive proteins or to supplement the patient's own production of PKIN. PKIN nucleic acids may be used to produce the PKIN polypeptide, by inserting the nucleic acids into a host cell and culturing the cell to express the protein. PKIN nucleic acid and its complementary sequences may also be used as DNA probes in diagnostic assays to detect and quantitate the presence of similar nucleic acid sequences in samples and therefore which patients may be in need of restorative therapy. The present sequence is human PKIN-9 cDNA.

Sequence 3141 BP; 685 A; 941 C; 942 G; 573 T; 0 other;

Alignment Scores:
 Pred. No.: 6,63e-40 Length: 3141
 Score: 445.00 Matches: 111
 Percent Similarity: 53.33% Conservative: 49
 Best Local Similarity: 37.00% Mismatches: 107
 Query Match: 27.55% Indels: 34
 DB: 22 Gaps: 8

US-09-830-144-2_COPY_1_303 (1-303) x AAD18824 (1-3141)

Qy 3 ThrAlSerAlaAlaSerSerSerSerSerAlaGlyGluMetIleGluAlaPro 22
 Db 334 ACCCGCGCGAGCCCTTTCACGCCCTGCCAGCCCGCGCGAG----- 378
 Qy 23 SerGlnValIleuAsnIleGluIleAspTyrLysGluIleGluValGluIleVal 42
 Db 379 -----GAGAAATTGATTTCGCGAGCTCACCCTTGGAAGAGATTATT 420
 Qy 43 GlyArgGlyAlaIleGlyValValCysLysAlaLysTyrArgAlaLysAspValAlaIle 62
 Db 421 GGCATCGGGGGCTTGGGAAGGCTCTATCGTCTTCTGATAGGGGATAGGTGCTGTG 480
 Qy 63 Lys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluIleuArgGln 80
 Db 481 AAAGCAGCTGGCCAGCCTTGATGAGCATGAGCCAGACCAATGAGAATGTTGCGCAA 540
 Qy 81 -----LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 96
 Db 541 GAGGCCAGCTTCGCCCATGCTGGAAGCACCACATCATTTGCCCTTAGAGGGGTATGT 600
 Qy 97 LeuAsnPro-----ValCysLeuValIleMetGlyTyrAlaGluGlySerLeuTyrAsn 114
 Db 601 CTGAAGAGGCCCACTTGCTGTGATGAGATTGCTCGTGGAGAGACCTTGAATAGA 660
 Qy 115 ValLeuHisGlyAlaGluProIleuProTyrTyrTyrThrAlaIleHisIleMetSerTyrCys 134
 Db 661 GTGTATTCTGGGAAAGAGATTCCTCCCAAGCATC-----CTGGTGAATTGGGCT 708
 Qy 135 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 154
 Db 709 GTCCAGATTGCCAGAGGAGTGAATCTTACTTGATGAGGCAATTGTCCTCATCTCCAC 768
 Qy 155 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 167
 Db 769 CGGACCTTAAGTCCAGCAACATATTGATCTCCAGAAAGTGGAGATGAGACCTGAGC 828
 Qy 168 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 186
 Db 829 AACCAAGATTCTGAAGATCTGATTGCTGGCTGGGAAATGGACCCAGCCACCAAG 888
 Qy 187 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheGluLysSerAsnTyr 206
 Db 889 ATGATGCGCGGAGGACGATATGCTTGATGGACCCGAAATGATCCGGGCTCCCATGTTT 948
 Qy 207 SerGlnLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArg 226
 Db 949 TCCAAAGGAGCATGATGTGGAGACTATGGGGTGCTACTTTGGGAGTTGCTGACTGGTAG 1008
 Qy 227 LysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAlaValHisAsnGly 246

Db 1009 GTGCCCTTCGAGGCAATTGATGGCTTAGCACT--CGC---TTATGAGATGGCCCATGACAA 1064
 Qy 247 ThrArgPro-ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCys 266
 Db 1065 ACTGCGCCTTCTTATCTCTTCTGATGCTGACGTCAGCAACCTTTGGCCAACTATGGAAGCTG 1124
 Qy 266 sTyrSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMetThr 285
 Db 1125 CTGGAATCTGATCCCATGACGACCATCTTTACAGAAATATCTGTGACCACTAAC 1182

RESULT 15

ID AAD34309 standard; cDNA; 3538 BP.
 AC AAD34309;
 DT 16-JUL-2002 (first entry)
 XX
 DE Human PKIN-12 cDNA.
 XX
 KW Human; kinase; enzyme; PKIN-12 protein; immune system disorder; anaemia; acquired immune deficiency syndrome; thymic hypoplasia; Crohn's disease; asthma; neurological disorder; epilepsy; Charcot-Marie-Tooth disease; AIDS; seizures; cell proliferative disorder; cancer; adenocarcinoma; leukaemia; lymphoma; melanoma; myeloma; sarcoma; developmental disorder; Down's syndrome; gene therapy; protein therapy; cyostatic; gene; ss.
 XX Homo sapiens.
 OS
 XX
 FH Key Location/Qualifiers
 FT CDS 1..3294
 FT /*tag= a
 FT /product= "Human PKIN-12 protein"
 FT s1g_peptide 1..51
 FT /*tag= b
 FT mat_peptide 52..3291
 FT /*tag= c
 FT /product= "Human mature PKIN-12 protein"
 FT
 PN WO200218557-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 31-AUG-2001; 2001WO-US27219.
 XX
 PR 31-AUG-2000; 2000US-229873P.
 PR 08-SEP-2000; 2000US-231357P.
 PR 14-SEP-2000; 2000US-232654P.
 PR 22-SEP-2000; 2000US-234902P.
 PR 29-SEP-2000; 2000US-236499P.
 PR 06-OCT-2000; 2000US-238369P.
 PR 13-OCT-2000; 2000US-240542P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Bandman O, Nguyen DB, Walla NK, Hafalia AUA, Yao MG, Gandhi AR, Guruxjan R, Ding L, Patterson C, Yue H, Baughn MR, Tribouley CM, Thornton M, Elliott JS, Lu Y, Jeon CH, Au-Young J, Tang YT, Azimzai Y, Burdill UD, Marcus GA, Zingler KA, Lu DM, Lai PG, Rankumar J, Warren BA, Kearney L, Policky JL, Thangavealu K, Burford N;
 XX
 DR WPI; 2002-329769/36.
 DR P-PSDB; AAE21717.
 XX
 PT New human kinases, useful for diagnosing, treating or preventing immune system disorders (e.g. Crohn's disease), neurological disorders (e.g. epilepsy), or cell proliferative disorders (e.g. cancers such as leukemia or lymphoma)
 PT
 PS Claim 91; Page 207-208; 218pp; English.
 XX
 CC The present invention relates to human kinases (PKIN) and polynucleotides

CC encoding such proteins. PKIN sequences of the invention are useful for
CC diagnosing, treating or preventing disorders associated with aberrant
CC expression of PKIN, particularly immune system disorders (e.g. acquired
CC immune deficiency syndrome (AIDS), thymic hypoplasia, Crohn's disease,
CC anaemia, asthma), neurological disorders (e.g. epilepsy, Charcot-Marie-
CC Tooth disease or seizures), cell proliferative disorders (e.g. cancers
CC such as adenocarcinoma, leukaemia, lymphoma, melanoma, sarcoma),
CC and developmental disorders (e.g. Down's syndrome). They are also used
CC in gene therapy and protein therapy. The present sequence is a cDNA
CC encoding human PKIN-12 protein.
XX

SQ Sequence 3538 BP; 763 A; 1055 C; 1062 G; 658 T; 0 other;

Alignment Scores:

Pred. No.:	1.34e-39	Length:	3538
Score:	443.00	Matches:	113
Percent Similarity:	53.97%	Conservative:	50
Best Local Similarity:	37.42%	Mismatches:	112
Query Match:	27.43%	Indels:	28
DB:	24	Gaps:	9

US-09-830-144-2_COPY_1_303 (1-303) x AAD34309 (1-3538)

QY	3	ThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGlu-----MetIleGlu	20
DB	334	ACCCGCGCAGCGCCTTCTCCAGCGCTGCCAGCGCGCGGCGGAGGACCCCACTTCTCTAC	393
QY	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrIlysGluIleGluValGluGlu	40
DB	394	CGGCCCATTCAGTTGTTA-----GAAATTGATTTTCGGAGCTCACCTTGGGAAGAG	444
QY	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal	60
DB	445	ATTATTGGCATCGGGGCTTTGGGAGGCTATCGTGCTTCGTGATAGGGATGAGGTT	504
QY	61	AlaIleLys-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeu	78
DB	505	GCTGTGAAGCAGCTCGCCACGACCTGTATGATGAGGACATCAGCCAGACCATAGAGAATGT	564
QY	79	ArgGln-----LeuSerArgValAsnHisProAsnIleValLysLeuTyrGly	94
DB	565	CGCCAGAGGCCAAGCTCTTCGCCATGCTGAAGACCCCAACATCATTCCTTAAGAGGG	624
QY	95	AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu	112
DB	625	GTATGCTGAAGGAGGCCAACCTCTGCTTGCTATGGAGTTTCTCTGTGGAGGACCTTTG	684
QY	113	TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer	132
DB	685	AATAGAGTGTATCTGGGAAAAGGATTCGCCACAGATC-----CTGGTGAAT	732
QY	133	TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu	152
DB	733	TGGCTGTGCAGATTGCGCAGGAGTGAATCTACTTGTATGAGGCAATTTGTCCCATC	792
QY	153	IleHisArgAspLeuLysProProAsnLeuLeuVal-----	165
DB	793	ATCACCCGCGACCTTAAGTCCAGCACATATTTGATCCTCCAGAGGTGGAGATGGAGAC	852
QY	166	AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr	184
DB	853	CTGAGCAACAAGATCTGAAGATCACTGATTTTGGCTCGGGAATGGCACCGCAACC	912
QY	185	HisMetThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPheGluGlySer	204
DB	913	ACCAAGATGATCGCGCAGGCGGTATGCTTGGATGGACCCCGAAGTCATCCGGGCTCC	972
QY	205	AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr	224
DB	973	ATGTTTCCAAAGGCAGTGATGTGGAGCTATGGGTGCTACTTTGGGAGTTGCTGACT	1032
QY	225	ArgArgLysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaValHis	244

Db	1033	GGTGAGGTGCCCTTTCGAGGCAATTGATGGCTTAGCACT-CGC--TTATGGAGTGGCCAT	1088
QY	245	AsnGlyThrArgPro-ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetTh	264
Db	1089	GAACAAACTCGCCCTTCTATTCCTTCTACGTGCCGAGAACCTTTTGCCAAACTCATGGA	1148
QY	264	rArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMe	284
Db	1149	AGACTGCTGGAATCCTGATCCCACTCAGCACCATCTTTCACGAATATCCTGGACCAGCT	1208
QY	284	tThr	285
Db	1209	AACC	1212

Search completed: December 10, 2002, 02:14:30

Job time : 341 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:04:15 ; Search time 2206 Seconds
(without alignments)
2224.493 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTASAASSSSSSSAGEMIE.....MTHLMRYFGADPELPVPCQ 303

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+_p2n.model -DEV=xlp
-Q=/cgn2_1/USP10_spo01/US09830144/runat_04122002_141511_2925/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plo -NORM=ext -HAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN_1_1_1716 @runat_04122002_141511_2925 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_nam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1379	85.4	1062	13	BM554120	BM554120 AGENCOURT
2	1242.5	76.9	998	9	AL525728	AL525728 AL525728
3	1167	72.3	739	13	BI696710	BI696710 603347214
4	1119	69.3	827	9	AL550389	AL550389 AL550389
5	1116	69.1	1006	14	BQ219348	BQ219348 AGENCOURT
6	1069	66.2	604	13	BM426610	BM426610 pgt2n.pk0
7	1063	65.8	929	9	AL520975	AL520975 AL520975
8	994.5	61.6	1054	14	BQ067693	BQ067693 AGENCOURT
9	984	60.9	646	13	BJ073883	BJ073883 BJ073883
10	982	60.8	688	13	BJ062988	BJ062988 BJ062988
11	982	60.8	696	13	BJ074867	BJ074867 BJ074867
12	953.5	59.0	910	12	BF780358	BF780358 602103276
13	921	57.0	747	13	BI334774	BI334774 602996777
14	920	57.0	542	10	AM960377	AM960377 EST372448
15	912	56.5	822	13	BI757369	BI757369 603025591
16	880.5	54.5	760	12	BG773488	BG773488 602720209
17	873	54.1	682	10	AV398933	AV398933 AV398933
18	852.5	52.8	745	13	BI767913	BI767913 603061057
19	843	52.2	681	10	AV398935	AV398935 AV398935
20	827.5	51.2	1191	13	EM472107	EM472107 AGENCOURT
21	823	51.0	641	14	BQ386875	BQ386875 NISC_mn20
22	814	50.4	1033	12	BE882182	BE882182 601505283
23	809	50.1	647	10	BB665645	BB665645 BB665645
24	794.5	49.2	594	13	BJ014293	BJ014293 BJ014293
25	748	46.3	686	12	BG548917	BG548917 602575750
26	741.5	45.9	585	9	AL773851	AL773851 AL773851
27	741	45.9	771	13	BI093821	BI093821 602860601
28	740.5	45.9	718	9	AL652289	AL652289 AL652289
29	720	44.6	516	13	BJ006787	BJ006787 BJ006787
30	715	44.3	586	13	BJ035129	BJ035129 BJ035129
31	714	44.2	605	9	AL704467	AL704467 DKF2p686A
32	712.5	44.1	634	9	AL710196	AL710196 DKF2p686A
33	712.5	44.1	700	10	BE291363	BE291363 601084896
34	704.5	43.6	831	14	BQ736206	BQ736206 AGENCOURT
35	689	42.7	845	12	BF027003	BF027003 601671216
36	681	42.2	642	10	BB645282	BB645282 BB645282
37	681	42.2	647	13	BJ008789	BJ008789 BJ008789
38	664.5	41.1	569	9	AUI77912	AUI77912 AUI77912
39	637	39.4	607	9	AIS41706	AIS41706 SD06739.5
40	609.5	37.7	644	9	AL782943	AL782943 AL782943
41	566.5	35.1	615	9	AL679013	AL679013 AL679013
42	557	34.5	858	13	BG916821	BG916821 602816079
43	553	34.2	711	10	AV718168	AV718168 AV718168
44	548	33.9	538	14	BQ557823	BQ557823 H4048E02-
45	532	32.9	447	9	AA306104	AA306104 EST177086

ALIGNMENTS

RESULT 1
BM554120
LOCUS
DEFINITION BM554120 1062 bp mRNA linear EST 20-FEB-2002
5', mRNA sequence.
ACCESSION BM554120
VERSION BM554120.1 GI:18793446
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lln.gov
 Plate: LLM12760 row: f column: 18
 High quality sequence stop: 730.
 Location/Qualifiers

FEATURES

source

1..1062
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5742353"
 /clone_1ib="NIH MGC 119"
 /tissue_type="medulla"
 /lab_host="DH10B"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: Not; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library."
 BASE COUNT 253 a 243 c 298 g 266 t 2 others
 ORIGIN

Alignment Scores:

Pred. No.: 7,71e-150 Length: 1062
 Score: 1379.00 Matches: 278
 Percent Similarity: 95.92% Conservative: 5
 Best Local Similarity: 94.56% Mismatches: 4
 Query Match: 85.39% Indels: 7
 DB: 13 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x BM554120 (1-1062)

OY 1 MetSerThAlSeSerAlaAlaSeSerSerSerSerSerAlaGlyGluMetIleGlu 20
 DB 183 AGTGTACAGCCTTCGCCCTTCCTCTCTCTCGCGCGGATGATCGAA 242
 OY 21 AlaProSerGlnValIleuAsnPhcGluGluIleAspTyrGlyGluIleGluValGluGlu 40
 DB 243 GCCCTTCCAGGCTCTCAACTTTGAAGAGATCGACTCAAGAGATCGAGGTGAAGAG 302
 OY 41 ValValGlyArgGlyAlaPhcGlyValValCysGlyAlaIleSTPArgAlaIleAspVal 60
 DB 303 GTTGTGGAGAGAGGCCCTTGGAGTGTGTCGAAGCTCAAGTGAAGAGCAAAAGATGTT 362
 OY 61 AlaIleGlyGlnIleGluSerGluSerGluSerGluSerGluSerGluSerGluSerGlu 80
 DB 363 GCTATTAAACAATAGAAAGATCTGAGAGAAAGCGTTATTGTAGAGCTTGGGAG 422
 OY 81 LeuSerArgValAsnHisProAsnIleValIleGluSerGlyValAlaCysLeuAsnProVal 100
 DB 423 TTATCCCGGTGAACCATCTTAATGTAAGCTTTATGAGCGCTGTTGATCCAGTG 482
 OY 101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValIleHisGlyValAlaGlu 120
 DB 483 TGTCTTGGATGGAATATGCTGAAGGGGCTTTTATATATGCTCAGTGGTGGTGA 542
 OY 121 ProLeuProTyrTyrThAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGlnGly 140
 DB 543 CCATTGCCATATTATACCTGCTCCCAAGCAATGATGTTGTGTTTACAGTGTTCACAGGA 602
 OY 141 ValAlaTyrLeuHisSerMetGlnProIleValAlaGluIleHisArgAspLeuIleProPro 160
 DB 603 GTGGCTTATCTTCAAGCATGCAACCAAGGGCTAATTCACAGGAGCCTGAACCAACA 662

OY 161 AsnLeuLeuValAlaGlyTyrValIleuLysIleCysAspPhcGlyThrAlaCys 180
 DB 663 AACTTACTGCTGGTGTGACAGGGGACAGTCTTAAATTTGATTTGGTACAGCCTGT 722
 OY 181 AspIleGlnThrHisMetTrrAsnAsnLysGlySerAlaAlaThrPheMetAlaProGluVal 200
 DB 723 GACATTGACACACATGACCAATACACAGGAGAGTGTCTTGATGGACCTGAAGTT 782
 OY 201 PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTyr 220
 DB 783 TTGAAGTAGTATATTACAGTGAATAATGACGCTTTCAGCTGGGGGINTATTCTTTGG 842
 OY 221 GluValIleThrArgArg-LysProPheAspGluIleGlyGly-ProAlaPheArgIleMet 240
 DB 843 GAAATGATAACGGCTCGGAACCCCTTTGATGAGATGGGTGGCCCAAGCTTCCGAATCA 902
 OY 240 eTTpAlaValHisAsnGlyThrArgProIleu-IleLysAsnLeuProLysProIle 259
 DB 903 TTGGGCTGCTCATATATGATGACTGACACCACTGATATATAATTTACCTAGGCCATT 962
 OY 260 Glu-SerLeuMetThr-ArgCysTrpSerLysAspPro-SerGlnArgProSerMetGlu 278
 DB 963 GAAAGCCTGATGATCCTCCTGNTTGGGCTTAAGATTTCTTCCAGGCGCTTTCATGAGAG 1022
 OY 279 -GluIleValIleLysIleMetThrHisLeu 287
 DB 1023 GGAATGTTGTAATAATGATGACTGCTTG 1050

RESULT 2

AL525728 998 bp mRNA linear EST 13-FEB-2001
 AL525728 LIT_NFL003_NBC3 Homo sapiens cDNA clone CS0DC013YE20 5
 LOCUS DEFINITION
 AL525728 prime, mRNA sequence.

ACCESSION AL525728
 VERSION AL525728.1 GI:12789221
 KEYWORDS EST.

SOURCE

human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 998)
 AUTHORS Li W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 CONTACT: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

COMMENT

FEATURES
 source
 1..998
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0DC013YE20"
 /clone_1ib="LIT_NFL003_NBC3"
 /sex="male"
 /tissue_type="neuroblastoma cells"
 /lab_host="DH10B"
 /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@life.com URL : http://fulllength.invitrogen.com"

BASE COUNT 245 a 225 c 268 g 247 t 13 others
 ORIGIN

Alignment Scores: 5.73e-134 Length: 998
 Pred. No.:

[illegible]

QY	21	AlaProSerGlnValAlaLeuAsnPhgIuGIuIleAspYrYrYsGluIleGIuValGIuIu	40
Db	217	GCCTTCCAGGCTCTCACTTTGAAGARITCGACITACAAAGAGATGAGCTGGAAAG	276
QY	41	ValValGIuYrGIuYalAlaPhgGIuValValCysIysAlaIalYrPrAlAlaLysAspVal	60
Db	277	GTTGTGGAGAGAGAGCTTTGGARTGTTTGGCAAGCTAAGTGAAGAGCAAAAATGT	336
QY	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	337	GCTATTAAATAAATAAAGTGAATCTGAGAGAAAGGCTTATTGTATGAGCTTGGAG	396
QY	81	LeuSerArgValAlaAsnHisProAsnIleValLysLeuYrGIuYalAlaCysLeuAsnProVal	100
Db	397	TTATCCGCTGTRAAACATCCATATATTGTAAAGCTTATGAGGCTGCTGATTCAGTG	456
QY	101	CysLeuValMetGluYrYrAlaGIuGIuGIySerLeuYrAsnValLeuHisGIuYalGlu	120
Db	457	TGCTCTTGAGAGGAATATGCTGAAGGGGGCTCTTATATTAATGTRCGACAGTGCTGA	516
QY	121	ProLeuProYrYrYrTrpAlaAlaHisAlaMetSerTrpYsLeuGlnCysSerGln-GI	140
Db	517	CCATTGCCAATATTAAGTCTGCCACGCAATGAGTGCTTTTACAGATGTTCCCAAG	576
QY	140	ValAlaIatYrLeuHisSerMetGlnProYsAlaLeuIleHisArgAspLeuYsProPr	160
Db	577	AGTGCTTATCTTACAGAGCATGCAACCCAAAGCCTTATTCACGGAGACTGAAACACC	636
QY	160	AsnLeuLeuLeuValAlaGIuGIyYrTrpValLeuLysIleCysAspPheGIuYrAlaC	180
Db	637	AAACTTACTGCTGGTGGAGGGGGAGACATCTTAAAAATTTGTATTTGGTTACAGCTG	696
QY	180	SAsIIleGlnTrhHisMetThrAsnAsnLysGIISeValAlaIleTrpMetAlaProGIuVa	200
Db	697	TGACATTCAGACACATGACCAATACAAAGGGAGTGCTCTTGAGTGGACCTGAAAGT	756
QY	200	IPheGIuGIySerAsnYrSerGIuLysCysAspValPheSerTrpGIuIleIleLeuTr	220
Db	757	TTTTGAAGGTGATATTACAGTAAAAATGAGCTCTTACGCTGGGGTATTATCTTTG	816
QY	220	pGIuValIle 223	
Db	817	GGAAGTGATA 826	
RESULT 5			
LOCUS	B0219348	1006 bp	mRNA linear EST 02-MAY-2002
DEFINITION	AGNCOCURT_7260864 NIH_MGC_71 Homo sapiens cdna clone IMAGE:5785294		
ACCESSION	B0219348		
VERSION	B0219348.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.		
TITLE	1 (bases 1 to 1006)		
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-r@mail.nih.gov		
	Tissue Procurement: ATCC		
	CDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNLN at:		
	http://image.llnl.gov		
	plate: LLAM12872 row: C column: 23		
	High quality sequence stop: 624.		
FEATURES	Location/Qualifiers		

QY	122	CACAGATGATAAACC	AAAGCCCTTAATTCACAGAGACCTG	AAACCAACCAATTTGCTTGC	181
QY <td>165 <td>VALAAGAGAGT <td>ThyValLeuLeuTylleCyAspPheGlyThyThalacysAsp11leGlnThr</td> <td>184</td> <td></td> </td></td>	165 <td>VALAAGAGAGT <td>ThyValLeuLeuTylleCyAspPheGlyThyThalacysAsp11leGlnThr</td> <td>184</td> <td></td> </td>	VALAAGAGAGT <td>ThyValLeuLeuTylleCyAspPheGlyThyThalacysAsp11leGlnThr</td> <td>184</td> <td></td>	ThyValLeuLeuTylleCyAspPheGlyThyThalacysAsp11leGlnThr	184	
Db <td>182 <td>GTAGCTGGGGGGAC <td>CTTAAAGATCTGTGATTTGGTACAGCCCTGTATATATTC <td>AAACA</td> <td>241</td> </td></td></td>	182 <td>GTAGCTGGGGGGAC <td>CTTAAAGATCTGTGATTTGGTACAGCCCTGTATATATTC <td>AAACA</td> <td>241</td> </td></td>	GTAGCTGGGGGGAC <td>CTTAAAGATCTGTGATTTGGTACAGCCCTGTATATATTC <td>AAACA</td> <td>241</td> </td>	CTTAAAGATCTGTGATTTGGTACAGCCCTGTATATATTC <td>AAACA</td> <td>241</td>	AAACA	241
QY <td>185 <td>HisMetThrAsnAsnLysGlySerAlaAla</td> <td>atPMeTalAProGluValPheGluGlySer</td> <td>204</td> <td></td> </td>	185 <td>HisMetThrAsnAsnLysGlySerAlaAla</td> <td>atPMeTalAProGluValPheGluGlySer</td> <td>204</td> <td></td>	HisMetThrAsnAsnLysGlySerAlaAla	atPMeTalAProGluValPheGluGlySer	204	
Db <td>242 <td>CACATGACCAACAA</td> <td>ATAAAGGAAGTGTCTGTGGATGGACACCTGAAGTTT</td> <td>TGAGAGTTAC</td> <td>301</td> </td>	242 <td>CACATGACCAACAA</td> <td>ATAAAGGAAGTGTCTGTGGATGGACACCTGAAGTTT</td> <td>TGAGAGTTAC</td> <td>301</td>	CACATGACCAACAA	ATAAAGGAAGTGTCTGTGGATGGACACCTGAAGTTT	TGAGAGTTAC	301
QY <td>205 <td>AsnTyrSerGluLysCysAspValPhe</td> <td>SerTrpGlyIleIleLeuTrpGluValIleThr</td> <td>224</td> <td></td> </td>	205 <td>AsnTyrSerGluLysCysAspValPhe</td> <td>SerTrpGlyIleIleLeuTrpGluValIleThr</td> <td>224</td> <td></td>	AsnTyrSerGluLysCysAspValPhe	SerTrpGlyIleIleLeuTrpGluValIleThr	224	
Db <td>302 <td>AATTACAGAGAAA</td> <td>AAAGTACGTTTCAGTTGGGGATTAATCTTTGGAGAGAAACAC</td> <td>361</td> <td></td> </td>	302 <td>AATTACAGAGAAA</td> <td>AAAGTACGTTTCAGTTGGGGATTAATCTTTGGAGAGAAACAC</td> <td>361</td> <td></td>	AATTACAGAGAAA	AAAGTACGTTTCAGTTGGGGATTAATCTTTGGAGAGAAACAC	361	
QY <td>225 <td>ArgAspGlysProPheAspGluIleGlyGlyProAla</td> <td>PheArgIleMetTrpValaValHis</td> <td>244</td> <td></td> </td>	225 <td>ArgAspGlysProPheAspGluIleGlyGlyProAla</td> <td>PheArgIleMetTrpValaValHis</td> <td>244</td> <td></td>	ArgAspGlysProPheAspGluIleGlyGlyProAla	PheArgIleMetTrpValaValHis	244	
Db <td>362 <td>CCTAGGAACCTTTT</td> <td>GCATATAGATGGTGTCCAGCTTTCGATATAATGTGGCAGTTTAC</td> <td>421</td> <td></td> </td>	362 <td>CCTAGGAACCTTTT</td> <td>GCATATAGATGGTGTCCAGCTTTCGATATAATGTGGCAGTTTAC</td> <td>421</td> <td></td>	CCTAGGAACCTTTT	GCATATAGATGGTGTCCAGCTTTCGATATAATGTGGCAGTTTAC	421	
QY <td>245 <td>AsnGlyThyArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThr</td> <td>264</td> <td></td> <td></td> </td>	245 <td>AsnGlyThyArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThr</td> <td>264</td> <td></td> <td></td>	AsnGlyThyArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThr	264		
Db <td>422 <td>AATGGTACTCGAC</td> <td>CCACCACTGATCAAAAACCTTAACCAATGAGAGTTTAATGACC</td> <td>481</td> <td></td> </td>	422 <td>AATGGTACTCGAC</td> <td>CCACCACTGATCAAAAACCTTAACCAATGAGAGTTTAATGACC</td> <td>481</td> <td></td>	AATGGTACTCGAC	CCACCACTGATCAAAAACCTTAACCAATGAGAGTTTAATGACC	481	
QY <td>265 <td>ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlnIleValLysIleMet</td> <td>284</td> <td></td> <td></td> </td>	265 <td>ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlnIleValLysIleMet</td> <td>284</td> <td></td> <td></td>	ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlnIleValLysIleMet	284		
Db <td>482 <td>CGCTGTGTGCTCA</td> <td>AGAGATCTCTACACCACTTCCATGAGAGAAATTTTAAATTAATG</td> <td>541</td> <td></td> </td>	482 <td>CGCTGTGTGCTCA</td> <td>AGAGATCTCTACACCACTTCCATGAGAGAAATTTTAAATTAATG</td> <td>541</td> <td></td>	CGCTGTGTGCTCA	AGAGATCTCTACACCACTTCCATGAGAGAAATTTTAAATTAATG	541	
QY <td>285 <td>ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys</td> <td>302</td> <td></td> <td></td> </td>	285 <td>ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys</td> <td>302</td> <td></td> <td></td>	ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys	302		
Db <td>542 <td>ACACACTTGATCG</td> <td>GGTATCTTCCNNNAGCTGATGAACCTCTCAGATATCTTGC</td> <td>595</td> <td></td> </td>	542 <td>ACACACTTGATCG</td> <td>GGTATCTTCCNNNAGCTGATGAACCTCTCAGATATCTTGC</td> <td>595</td> <td></td>	ACACACTTGATCG	GGTATCTTCCNNNAGCTGATGAACCTCTCAGATATCTTGC	595	
RESULT 7					
AL520975					
LOCUS	AL520975	LTI_NFL004_NBC2	Homo sapiens cDNA clone CS0DB002YM17	5	
DEFINITION	AL520975	prime, mRNA sequence.			
ACCESSION	AL520975				
VERSION	AL520975.1	GI:12784468			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Li W.B., Gruber C., Jeesee J. and Polayes D.				
TITLE	Full-length cDNA libraries and normalization				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: segrel@genoscope.cns.fr, Web : www.genoscope.cns.fr.				
FEATURES					
Source					
	1..929				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="CS0DB002YM17"				
	/clone_id="LTI_NFL004_NBC2"				
	/sex="male"				
	/tissue_type="neuroblastoma cells"				
	/lab_host="DH10B"				
	/note="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-Oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"				
BASE COUNT	213 a	235 c	249 g	222 t	10 others
ALIGNMENT SCORES:					

Pred. No.:	4,33e-113	Length:	929
Score:	1063.00	Matches:	204
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.51%	Mismatches:	0
Query Match:	65.82%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-Z_COPY_1_303 (1-303) x AL520975 (1-929)	
QY 1 MetSerThrAlaSerAlaIaISerSerSerSerSerSeraIaGIyMeIIeglu 20	
Db 305 ATGTCTACAGCCCTCGCGCCTCTCTCTGTCCTTGCGGCCGGATGATGCAA 364	
QY 21 AlaProSerGlnValLeuAsnPhenylLugluIlleapryrLySGluIIegluValGIglu 40	
Db 365 GCCCTTCCCAAGTCTCACTTTGAAMAATCACTAACAGAGATCGAGTGGAMAR 424	
QY 41 ValValGIyArgGlyAlaPheGIyValValCySySaIalysTPArXAlaIyaspVal 60	
Db 425 GTTGTGGAAGAAGAGACTTTGGAGTTCTTTGCCAAGCTAAGTAGAGACAAAATACTT 484	
QY 61 AlaIleySGinIIegluSerGIuSerGIuArgIySaIaphelIevaIGluLeuArgIn 80	
Db 485 GCTATTAAACAATAAGAAAGTAATCTGAGAGGAAGCGTTATTATTGAGCTTGGCAG 544	
QY 81 LeuSerArgValaenHisProasnIIEvalIySeurTyrgIyAlaCySleuAsnProVal 100	
Db 545 TTATCCCGTGGAAACCATCTTAATTGTAAAGCTTTATGAGCGCTGTTGAATCCAGTG 604	
QY 101 CysLeuValmetGIuTyraIaGIUGlYGlySerIeuTyraEnaVlleuHisGIyAlaGIu 120	
Db 605 TGTCTTGGAAGGAATATGCTGGAAGGGGGCTTTATATATATGTCGTGCATGTGCTGAA 664	
QY 121 ProIeuProTyrrTyThraIalaHisIaIawetSerTpCysIeuGIuCySSerGIuGIy 140	
Db 665 CCATTGCCATATTATACTGCTGCCACGCAATGAGTTGGTGTTCACAGTGTTCACAGGA 724	
QY 141 ValAlaTyrlLeuHISserMetGlnProIySaIaleuIleHisIaIgaAPLeuIySPro 160	
Db 725 GTGGCTTATCTTCAACAGATGCCAACCAAAGCGCTTAATTCACAGAGGACCTGAAACCACA 784	
QY 161 AsnLeuLeuLeuValAlaGIyGIyThrValIeuIySIeCyasAPPheGIyThraIaCys 180	
Db 785 AACTTACTGCTGGTTCAGAGGGGACAGTTCTTAAAAATTTGTGATTTTGGTACAGCCTGT 844	
QY 181 AspIlegInthrIaIwetThrAsnaSnLIySGIySerAlaIatIpMetAlaProGIuVal 200	
Db 845 GACATTGCAGACACATGACCAATPAACAAGGGAGTCTGCTTGGATGCMCCTGAAATT 904	
QY 201 PheGIUgIySerAen 205	
Db 905 TTGAAAGTATGTAAT 919	
RESULT 8	
LOCUS B0067693 1054 bp mRNA linear EST 02-APR-2002	
DEFINITION ASENOCOURT_6643018 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5768401	
ACCESSION B0067693	
VERSION B0067693.1 GI:19696739	
KEYWORDS EST.	
SOURCE human.	
ORGANISM Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.	
REFERENCE NIH-MGC http://mgc.nci.nih.gov/	
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)	
TITLE Unpublished (1999)	
JOURNAL Contact: Robert Strausberg, Ph.D.	
COMMENT Email: cgabs-r@mail.nih.gov	
CDNA Library Preparation: Life Technologies, Inc.	
tissue Procurement: Life Technologies, Inc.	

QY	134	CysLeuGlnCysSerGlnGlyValAlaTyrLeuHISerMetClnProCysAlaLeuLeu	153
Db	63	TGTTTACAATGTGGCCCAAGAGAGTTCATATTACATACATGACATGAAGCCAAAGGCTTGATT	122
QY	154	HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle	173
Db	123	CACAGGGAGCTCCAAACCCAACTTGTTGCTGTGTAGCTGAGGACATGTTCTTAAAGATT	182
QY	174	CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla	193
Db	183	TGTGACTTGTGTACGCTGTGATTCATTCACATCTCAATGACTTAATPACAAAGAAAGTCA	242
QY	194	AlaIrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe	213
Db	243	GCATGATGAGCTCCAGAAAGTTTTGAAGGTAGCACTACAGCGAAAATGATGACGTGTTT	302
QY	214	SerTrpGlyIleIleLeuTyrPglValIleThrArgArgLysProPheAspGluIleGly	233
Db	303	AGTTGGGGCATTAATTCCTTTGGGAAAGTAATACCCGAAAGAAACCTTTCGATGAAATTTGCT	362
QY	234	GlyProAlaPheArgIleMetTyrPalaValHisAsnGlyThrArgProProLeuIleLys	253
Db	363	GGTCACAGGTTCCGTATATATGTTGGGCGTTCACAAAGTACTCGGCCACCATTAATTAA	422
QY	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
Db	423	AATTTGCTTAAGCTATTGAAAGCTTAAGCAGCTGCTGCTGCCAAAGATCCCCCAA	482
QY	274	ArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGly	293
Db	483	AGACCTTCAAATGAGGAGGAGATTGTCMAAGATATGACACATTAACAGATATTTCTGGA	542
QY	294	AlaAspGluProLeuGlnTyrProCysGln	303
Db	543	GCAGACGTTTCTTACAGATTCCTTCTCAG	572
RESULT 10			
LOCUS	BJ062988	688 bp	mRNA linear EST 10-DEC-2003
DEFINITION	BJ062988 NIBB Mochii normalized Xenopus tailbud library Xenopus		
ACCESSION	BJ062988		
VERSION	BJ062988.1	GI:17470746	
KEYWORDS	EST.		
SOURCE	African clawed frog.		
ORGANISM	Xenopus laevis		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
	Xenopodidae; Xenopus.		
REFERENCE	1. (bases 1 to 688)		
AUTHORS	Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and Kohara		
	Y.		
TITLE	Expressed genes in X. laevis embryo		
JOURNAL	Unpublished (2001)		
COMMENT	Contact: Tadaeu Shin-i		
	Center For Genetic Resource Information		
	National Institute of Genetics		
	1111 Yata, Mishima, Shizuoka 411-8540, Japan		
	Tel: 81-559-81-6856		
	Fax: 81-559-81-6855		
	Email: tshini@genes.nig.ac.jp.		
FEATURES			
source			
	1..688		
	/organism="Xenopus laevis"		
	/db_xref="taxon:8155"		
	/clone="XL069m17"		
	/clone_lib="NIBB Mochii normalized Xenopus tailbud		
	library"		
	/tissue_type="whole embryo"		
	/dev_stage="stage 25"		
	/note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs		
	were oligo-dT primed and directionally cloned. Staging		
	according to Nieuwkoop and Faber. Library is subtracted		

BASE COUNT	206 a	139 c	154 g	168 t	1 others
ORIGIN	(Mellicome/CRIC Institute). "				
Alignment Scores:					
Pred. No.:	7,45e-104				
Score:	982.00				
Percent Similarity:	97.83%				
Best Local Similarity:	96.20%				
Query Match:	60,80%				
DB:	13				
	Gaps: 0				
US-09-830-144-2_COPY_1_303 (1-303) x BU062988 (1-688)					
Qy 120	GIUPProleupProTyrTrpThraAlaHisAlaMetSerTPCySleuGlnCysSerGln 139				
Db 22	GACCTTGGCTTACCTACCTACTGCTGCTGCCAATGCAAGAGTGGTGTTCACATGGGCCAA 81				
Qy 140	GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 159				
Db 82	GGAGTGGATATTATTAACATAGACGAAGCCAAAGGCTGTGATTCACAGGAGCTCAACCA 141				
Qy 160	ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 179				
Db 142	CCAACTGTTGGCTGGAGCTGGAGGCACTGTTCTTAAGATTGGACTTGGTACAGCC 201				
Qy 180	CysAspIleGlnThrHisMetThrAsnLeuLysGlySerAlaAlaTyrMetAlaProGlu 199				
Db 202	TGTGATTTGAGCTCAGCTCAGCTACTAATTAACAAAGAAATGACAGATGATGGCTCCAGAA 261				
Qy 200	ValPheGlnGlySerAsnTyrSerGlnLysCysAspValPheSerTyrGlyIleIleLeu 219				
Db 262	GTTTGTGAAGTGAAGCACTACAGCCGAAATGTGACGTTTATGTGGGCAATTATCTT 321				
Qy 220	TyrGluValIleThrArgArgLysProPheAspGlnIleGlyGlyProAlaPheArgIle 239				
Db 322	TGGGAATTAATACCCGAAGAAACCTTTCAGTGAATGGTGGTCCAGGGTCCGTAA 381				
Qy 240	MetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259				
Db 382	ATGGGCGCTGTTCACATGATGTACTGGGCAACATTAAATTAATTTGCTTAAGCCATT 441				
Qy 260	GlnSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgProSerMetGlnGlu 279				
Db 442	GAAAGCTTAATGCTGCTGCTGGTCCAAAGATCCCCCAAAACCTTCAATGGAGGAG 501				
Qy 280	IleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 299				
Db 502	ATTGTCAAGTATATGACATCTTAAGCAGTATTTTCTGGAGCAGACGTTTCTTAACAG 561				
Qy 300	TyrProCysGln 303				
Db 562	TATCCTGTGTCAG 573				
RESULT 11					
LOCUS	BU074867 696 bp mRNA linear EST 11-DEC-2001				
DEFINITION	BU074867 NIBB Mochi normalized Xenopus laevis caildry library Xenopus laevis cDNA clone XL071110 5', mRNA sequence.				
ACCESSION	BU074867				
VERSION	BU074867.1 GI:17505056				
KEYWORDS	EST.				
SOURCE	African clawed frog.				
ORGANISM	Xenopus laevis				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus. 1 (bases 1 to 696).				
AUTHORS	Kityama,A., Terasaka,C., Mochi,M., Ueno,N., Shin-i,T. and Kohara Y.				
TITLE	Expressed genes in X. laevis embryo				
JOURNAL	Unpublished (2001)				

COMMENT

Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers

FEATURES

source
1..696
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="X1071110"
/clone_lib="NIBB Mochii normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 210 a 142 c 155 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 7.59e-104 Length: 696
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 60.80% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x BJ074867 (1-696)

QY 120 GluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 139

Db 22 GAACCTTTGGCTTACTATACGTGGCCATGCAATGAGTTGGTTTACAAATGGCCAA 81

QY 140 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 159

Db 82 GGAGTTGCATATTACATCATCATGAGCCAAAGGCTCTGATTCACAGGACCTCAACCA 141

QY 160 ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 179

Db 142 CCAAACTTTGTGTGTAGCTGGAGGCACGTCTTAAAGATTGTGACTTTGGTACAGCC 201

QY 180 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIaIaIaIaIaIaIaIaIa 199

Db 202 TGTGATATTACAGACTACATGACTTAATACAAAGGAAGTGCAGCATGGATGGCTCCAGAA 261

QY 200 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeu 219

Db 262 GTTTTTCGAGGTAGCAACTACAGCAAAATGTGACGTGTTTGTAGTTGGGCATATTCTT 321

QY 220 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 239

Db 322 TGGGAAGTAATAACCCGAAGAAACCTTTTCGATGAAATTTGGTCCAGCGTTCCGTATA 381

QY 240 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 259

Db 382 ATGTGGGCTGTTCACATGGTACTCGGCCACCATTAATTAATAAATTTGGCTTAAGCCTATT 441

QY 260 GluSerLeuMetThrArgCysTTrpSerLysAspProSerGlnArgProSerMetGluGlu 279

Db 442 GAAGCTTAATGACTCGCTGCTGGTCCAAAGATCCCCCAAGACCTTCAATGGAGGAG 501

QY 280 IleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 299

Db 502 ATGTCAAGATAATGACATCTAAAGCAGTAATTTCTTCGAGCAGACGTTTCCTTACAG 561

QY 300 TyrProCysGln 303

Db 562 TATCCTTGTTCAG 573

RESULT 12

BF780358

LOCUS

DEFINITION

ACCSSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

1 (Bases 1 to 910)

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Jeffrey E. Green, M.D.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LENL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LENL at:

http://image.llnl.gov

Plate: LLAM9806 row: h column: 20

High quality sequence stop: 664.

Location/Qualifiers

source

1..910

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:4221379"

/clone_lib="NCI CGAP Kid14"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI;

Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.

Average insert size 1.75 kb. Constructed by Life

Technologies. Note: this is a NCI CGAP Library. |"

BASE COUNT 247 a 209 c 246 g 207 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 2.45e-100 Length: 910
Score: 953.50 Matches: 194
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 92.38% Mismatches: 6
Query Match: 59.04% Indels: 12
DB: 12 Gaps: 1

US-09-830-144-2_COPY_1_303 (1-303) x BF780358 (1-910)

QY 103 ValMetGluTyrAlaGluGlyGlySer-Leu-----TyrAsnVa 115

Db 4 GTGATGGAATATGCAGAGGGGGCTCATGTAATATGTTTTCGTGCCTTCTCTTACAGT 63

QY 115 IleHisGlyValAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLe 135

Db 64 GCTGCATGGTGTGAACCATTTGCCCTTACTACACTGCTCTCATGCCATGAGCTGGTGT 123

QY 135 uGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisAr 155

Db 124 ACAGTGTTCACAGAGTGGCTTACCTGCACAGCATGCAGCCCAAGGCTGATTCACAG 183

QY 155 gAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAs 175

Db 184 GGACCTCAAGCTCCAAACTTGTCTGCTGGTGCAGAGGGACAGTCTTAAAAATCTGCGA 243

QY 175 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 195

Db 244 TTTTGTGACGCTTGTGATCCTCAACACACATGACCAATAATAAGGGAGTGCTGCTTG 303

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:07:24 ; Search time 71 seconds
(without alignments)

1308.775 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTAASASSSSSSSAGEMIE.....MTHLMRYFGADEPLQVPCQ 303

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 1533381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141511_2942/app_query.fasta_1.455
-DB=Issued Patents NA -QWTF=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPTXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144@cgn2_1_1_31@runat_04122002_141511_2942 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUT.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1615	100.0	1788	US-09-529-279-14	Sequence 14, Appl
2	1615	100.0	2656	US-08-685-625A-5	Sequence 5, Appl
3	1615	100.0	2856	US-09-529-279-3	Sequence 3, Appl
4	1604	99.3	2443	US-08-685-625A-1	Sequence 1, Appl
5	442	27.4	2120	US-09-221-235-4	Sequence 4, Appl
6	442	27.4	2120	US-09-221-928-4	Sequence 4, Appl
7	442	27.4	2120	US-09-221-527-4	Sequence 4, Appl
8	442	27.4	2120	US-09-221-236-4	Sequence 4, Appl
9	442	27.4	2120	US-09-221-416-4	Sequence 4, Appl
10	442	27.4	2120	US-09-221-245-4	Sequence 4, Appl
11	442	27.4	2120	US-09-163-115-4	Sequence 4, Appl
12	442	27.4	2120	US-09-221-528-4	Sequence 4, Appl

13	442	27.4	2120	4	US-09-593-553-4	Sequence 4, Appl
14	442	27.4	2120	4	US-09-221-237-4	Sequence 4, Appl
15	433	26.8	1365	3	US-09-221-235-6	Sequence 6, Appl
16	433	26.8	1365	3	US-09-221-928-6	Sequence 6, Appl
17	433	26.8	1365	3	US-09-221-527-6	Sequence 6, Appl
18	433	26.8	1365	3	US-09-221-236-6	Sequence 6, Appl
19	433	26.8	1365	3	US-09-221-416-6	Sequence 6, Appl
20	433	26.8	1365	4	US-09-221-245-6	Sequence 6, Appl
21	433	26.8	1365	4	US-09-163-115-6	Sequence 6, Appl
22	433	26.8	1365	4	US-09-221-528-6	Sequence 6, Appl
23	433	26.8	1365	4	US-09-593-553-6	Sequence 6, Appl
24	433	26.8	1365	4	US-09-221-237-6	Sequence 6, Appl
25	410	25.4	2505	4	US-09-291-839-3	Sequence 3, Appl
26	410	25.4	3025	4	US-09-291-839-1	Sequence 1, Appl
27	401	24.8	3389	1	US-08-395-580-1	Sequence 1, Appl
28	401	24.8	3426	1	US-08-205-018-1	Sequence 1, Appl
29	398	24.6	2890	1	US-07-928-464-1	Sequence 1, Appl
30	398	24.6	2890	5	PCT-US93-07347-1	Sequence 1, Appl
31	398	24.6	3033	1	US-08-003-311B-1	Sequence 1, Appl
32	398	24.6	3033	1	US-08-261-432-1	Sequence 1, Appl
33	357.5	22.1	2574	4	US-09-142-529-2	Sequence 2, Appl
34	352	21.8	4508	5	PCT-US93-06251-34	Sequence 34, Appl
35	350	21.7	3132	1	US-08-167-919A-9	Sequence 9, Appl
36	350	21.7	3132	3	US-08-715-106-9	Sequence 9, Appl
37	344.5	21.3	3254	1	US-08-162-809-15	Sequence 15, Appl
38	339.5	21.0	3516	3	US-09-188-930-257	Sequence 257, App
39	338.5	21.0	4304	4	US-08-368-776A-1	Sequence 1, Appl
40	338.5	21.0	4304	5	PCT-US96-00419-1	Sequence 1, Appl
41	333.5	20.7	4529	2	US-08-449-645A-16	Sequence 16, Appl
42	333.5	20.7	4529	2	US-08-702-367A-16	Sequence 16, Appl
43	333.5	20.7	4529	5	PCT-US95-04681-16	Sequence 16, Appl
44	333	20.6	1398	2	US-08-604-989A-9	Sequence 9, Appl
45	333	20.6	1521	2	US-08-604-989A-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-529-279-14
; Sequence 14, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:
Pred. No.: 9.7e-187
Score: 1615.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 4
Length: 1788
Matches: 303
Conservative: 0
Mismatches: 0
Indels: 0
Gaps: 0

US-09-830-144-2_COPY_1_303 (1-303) x US-09-529-279-14 (1-1788)

Db	483	TGCTCTTGATGGAATATGCTGAGGGGGCTCTTTATATATAATGTGCTGCATGTGCTGAA	542
Qy	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	543	CCATTGCCATATTATACTGCTGCCACGCAATGAGTTGGTGTTTACAGTGTTCCTCAAGGA	602
Qy	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuLleHisArgAspLeuLysProPro	160
Db	603	GTGGCTTATCTTCACAGCATGCACCAAGCGCTAAATTCACAGGGACCTCGAAACCA	662
Qy	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysLleCysAspPheGlyThrAlaCys	180
Db	663	AACCTTACTGCTGGTTCGAGGGGGGACAGTTCTTAAAAATTTGTGATTTTGGTACAGCCTGT	722
Qy	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	723	GACATTTCAGACACACATGACCAATAACAAGGGAGTGTCTTGGATGGCACCTCGAAGTT	782
Qy	201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyLleIleLeuTrp	220
Db	783	TTTGAAGGTAGTAATTACAGTGAAAAAATGTGACGCTCTTCAGCTGGGGTATTATTCTTTGG	842
Qy	221	GluValIleThrArgArgLysProPheAspGluLleGlyGlyProAlaPheArgIleMet	240
Db	843	GAACTGATAACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCACAGCTTCCGAATCATG	902
Qy	241	TrpAlaValHisAsnGlyThrArgProProLeuLleLysAsnLeuProLysProIleGlu	260
Db	903	TGGGCTGTTTCATAATGGTACTCGACCACCACTGATAAAAAATTTTACCTAAGCCCATTTGAG	962
Qy	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	963	AGCCTGATGACTCGTTGGTCTAAGAATCCCTCCAGCGCCCTTCATATGGAGAAATT	1022
Qy	281	ValLysIleMetThrHisLeuMetArgTyrPheProGlyValAspGluProLeuGlnTyr	300
Db	1023	GTGAAAAATAATGACTCACCTTGATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT	1082
Qy	301	ProCysGln	303
Db	1083	CTTTGTGTCAG	1091

```

RESULT 3
US-09-529-279-3
; Sequence 3, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-09-529-279-3

```

Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0
US-09-830-144-2_COPY_1_303 (1-303) x US-09-529-279-3 (1-2656)			
Qy	1	MetSerThrAlaSerAlaAlaSerSerSerSerSerSerSerAlaGlyGluMetIleGlu	20
Db	183	ATGTCTACAGCCTCTCGCGCCCTCCCTCCTCCTCGTCTTCGGCGGFGAGATGATCGAA	242
Qy	21	AlaProSerGlnValLeuAsnPheGluGluIleAspTyrLysGluIleGluValIleGlu	40
Db	243	GCCCCITCCCAGGTCTCAACTTTGAGAGATCGACTACAAGGAGATCGAGTGGAAAGAG	302
Qy	41	ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal	60
Db	303	GTTGTTGGAAGAGAGCCCTTGGAGTTGTTTGCAAAGCTAAGTGGAGAGCAAAAAGATGTT	362
Qy	61	AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleValIleGluLeuArgGln	80
Db	363	GCTATTAAACAATAAGAAAGTGAATCTGAGAGAAAGCGTTTATTGTACAGCTTCGGCAG	422
Qy	81	LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal	100
Db	423	TTATCCCGTGTGAACCATCCTAATATTGTAAAGCTTTATGAGCGCTGCTTGAATCCAGTG	482
Qy	101	CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	483	TGTCCTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTGCTCGATGGTGCTGAA	542
Qy	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	543	CCATTGGCATATTATACTGCTGCCACGCAATGAGTTGGTGTTACAGTGTTCACAGGA	602
Qy	141	ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro	160
Db	603	GTGGCTTATCTTCACAGCATGCAACCAAGCGCTAAATTCACAGGGAGCTGAAACCA	662
Qy	161	AsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	663	AACTTACTGCTGTGTCAGGGGGACAGTTCTAAAAAATTTGTGATTTTGGTACAGCCTGT	722
Qy	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal	200
Db	723	GACATTGAGACACACATGACCAATAACAGGGGAGTGCTGTCTGATGGCACCTGAAGTT	782
Qy	201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp	220
Db	783	TTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTTCAGCTGGGGTATTATTCTTTGG	842
Qy	221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	843	GAAGTGATAACGGCTCGGAACCCCTTTGATGAGATTGGTGGCCCAAGCTTTCGAATCATG	902
Qy	241	TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	903	TGGGCTGTTCATAATGTACTCGACCAACCATGATAAAAAATTTTACCTAAGCCCATTCAG	962
Qy	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	963	AGCCTGATGACTCGTTGTGGTCTAAAGATTCCTTCCACGCGCCCTTCAATGGAGAAATT	102
Qy	281	ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr	300
Db	1023	GTGAAATAATGACTCACTTGATGCGGTACTTTCCAGAGGAGATGAGCCATTACAGTAT	108
Qy	301	ProCysGln	303
Db	1083	CTTTGTGAC	1091

RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A

Patent No. 5945301
GENERAL INFORMATION:
APPLICANT: UENO, Naoto
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: IRIE, Kenji
TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
TITLE OF INVENTION: TRANSDUCTION SYSTEM
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/685,625A
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-253549
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 001560-267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2443 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 157..1893
US-08-685-625A-1

Alignment Scores:		Pred. No.:		Length:	
Score:	Percent Similarity:	Score:	Percent Similarity:	Matches:	Conservative:
Best Local Similarity:	Query Match:	Best Local Similarity:	Query Match:	Mismatches:	Indels:
DB:		DB:		Gaps:	
US-09-830-144-2_COPY_1_303 (1-303) x US-08-685-625A-1 (1-2443)					
QY 1 MeSeerThraLaSeRaLaLaSeSeSeSeSeSeSeSeSeSeRaLaGluGluNeTtleGlu 20					
Dp 157 ATGTCGACAGCCCTCCGCCCTCGCTCCCTCCCTCGCTTCTCGCAGTAGATGCAA 216					
QY 21 AlaProSeRcInValLeuAnshheGluLuleAspYrLySgluLleGluValGluGlu 40					
Dp 217 GCCCCGTCGAGAGTCTGTGAACCTTCGAAAGATGCACACAGAGATGCAGGTGGAAAG 276					
QY 41 ValValG1AArgG1ValaIaPheG1ValVala1CysAlaYS1AlaYrYrPaRg1AlaYSaPVal 60					
Dp 277 GTTGTCGGAAGAGACCTTTTGAGATGATTTCGAAACCTTAAGTCGAGACCAAAAGATTTC 336					
QY 61 Ala1IeLysGln1IeGluSeRcIuSeRcIuAArg1YsAlaPhe1IeValGluLeuAArg1n 80					
Dp 337 GCATTTAACAGATGAAAGTAGTGCAGTCGAGGAGGCTTCTATGTGGAGCTCCGCGAG 396					
QY 81 LeuSeRarGValaShhIsProKsn1IeValLysLeuTyG1ValaCysLeuAnProVal 100					

Db	397	TTTGCGCGTGTGAACACATCTCTMAATTTGTCAAGTTGACGGAGCCCTGCCGATCCAGTA	456
OY	101	CysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGlyAlaGlu	120
Db	457	TGCTTTGATGAGGAATATGCAAGGGGGGGCTCATTTGATATATGTGTCGACATGGTCTGAA	516
OY	121	ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly	140
Db	517	CCATTGCCCTTACTACACTGCTGCTCCATGAGCCATAGAGTGCGTTTAAAGTGTCCCAAGA	576
OY	141	ValAlaTyrLeuHisSerMetGlnProLeuAlaLeuLHisArgAspLeuLysProPro	160
Db	577	GTGGCTTACCTGCACACACATGCACCCAAAGCCGTGATTACAGGGAGCCTCAAGCTCCA	636
OY	161	AsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCys	180
Db	637	AACTTGCTGCTGATTGCAGAGAGGACAGTCTTAAATACTGCATTTGGTACAGCTTGT	696
OY	181	AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaThrMetAlaProGluVal	200
Db	697	GACATCCAAACACACATGACCATAATAAAGGAGAGTGCCTTGATGGGGCTCGTAAGTG	756
OY	201	PheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp	220
Db	757	TTTGAAGGTACCAATTACAGTGAAGAAATGTGATGCTTCAAGCTGGGGGTATTATCCTCTGG	816
OY	221	GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet	240
Db	817	GAACTGTATACACCCCGGAAACCTTGATGAGATCGGTGGCCACATCTTCAGATCATG	876
OY	241	TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu	260
Db	877	TGGCGTGTCTAATATGGCACTTCGACCACTATCAAAAATTATTAAGCCCATTTGAG	936
OY	261	SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle	280
Db	937	AGCTTGATGACACCGCTGTGTCTAAGAGACCACTTCAGGGCCCTTCATATGAGGAAATT	996
OY	281	ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAlaAspGluProLeuGlnTyr	300
Db	997	GTGAAATTAATGACTCACTGATGGGTACTTCCAGAGAGCGGATGAGCATTAACGATAT	1056
OY	301	ProCys 302	
Db	1057	CCTTGT 1062	

```

US-09-221-235-4
; Sequence 4, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ. ID NOS.: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-235-4

Alignment Scores:
Pred. No.: 1.02e-43
Score: 442.00
Length: 2120
Matches: 105

```



```

Db      620 TATGCTGTGGTCTCTGGGAGATCTTAACAAGGAGGTCCTCCCTTTAAAGCTTTGGAAGA 679
Qy      235 ProAlaPheArgIleMetTrp--AlaValHisaenGlyThrArgProProIleuIleLys 253
Db      680 -----TTACAAGTGTAGCTTGCGCTGTACTGTGAAAAAAGAGAGATTACCATTTCCAGC 733
Qy      254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db      734 AGTTGCCCCAGAAAGTTTGCTGGAAGTGTACTGATCAGTGTGTGGAAAGCTGATGCCAAGAA 793
Qy      274 ArgProSerMetGluGluIleValLysIleMet 284
Db      794 CGGCCATCATTCACAAATCATTTCAATCTCG 826

RESULT 7
US-09-221-527-4
; Sequence 4, Application US/09221527
; Patent No. 6146832
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MN-050
; CURRENT APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-12-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-527-4

Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 3 Gaps: 12

US-09-830-144-2_COPY 1_303 (1-303) x US-09-221-527-4 (1-2120)
Qy      10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
Db      2 TCGACCCACGGCGTCGGGAGAAATATATATCTTTGTCATTATGAGATGTCGTCTCGGT 61
Qy      26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly 45
Db      62 GCCCTCTTTTGCAAAATTAATTTGATGACCTTCGACGTTTTTGAAAACTCGGAGGAGA 121
Qy      46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
Db      122 AGTTTGGGAAGTGTATTATCGAGCCAAATGCAATTCACAGACAGCAAGAGGTGGCTGTAAAG 181
Qy      64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db      182 AAGCTCTCTCAAAATAGAGAAAGAGGACGAA-----ATA 214
Qy      81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db      215 CTCAGTGTCTTCAGTTCACAGAAACATCATCCAGTTTATGAGATTAATTCCTGAACTCC 274
Qy      100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLysHisGly 118
Db      275 AACATGAGCATGTGTACAGAAATATAGCTTCTGTGGAGTCACTGTATGATTATTAACAGT 334
Qy      119 -----AlaGluProLeuProTyrTyrThrAlaIleHisAlaMetSerTyrCysLeuGln 136

```

```

Db      335 AACGAAAGTGAAGGAGT-----GATATGATCATCATTTAGACTGGCCACATGAT 385
QY      137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProlysalaleuLleHisArgAsp 156
      ::::::::::::::::::::
Db      386 GTAGCCAAAGGAATGCATTATTTCATATGAGAGGCTCTGTCAAGGTGATTCACAGAGAC 445
QY      157 LeuIysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuIysIleCysAspPhe 176
      ::::::::::::::::::::
Db      446 CTCAGGTCAAAAGAACGTTGTTATGCGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY      177 GlyThrAlaCysAspLleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
      ::::::::::::::::::::
Db      503 GGT---GCCTTCGGTTCATATCAATCAACAACACACATGCTCTGTTGGAGACTTTCOCA 559
QY      195 TrpMetAlaProGluValAlaPheGlnGlySerAsnLysSerGlnLysCysAspValAlaPheSer 214
      ::::::::::::::::::::
Db      560 TGGATGGCTCCAGAAAGTTATCCAAAGCTCCCTGTGTCAAAACTTGTGCACATATTC 619
QY      215 TrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGlnIleGly 234
      ::::::::::::::::::::
Db      620 TATGTTGGTGTCTCTCGGAGATCTTAACAAGGAGAGCTCCCTTTAAAGTTGGAAAGA 679
QY      235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyLysArgProProLeuLleLys 253
      ::::::::::::::::::::
Db      680 -----TTCAAGTAGAGCTTGCTTGTGTGTGTGAGAAAAAACAAGAGATTAAACCATTTCCAAGC 733
QY      254 AsnLeuProLysProIleGlnSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
      ::::::::::::::::::::
Db      734 AGTTGCCCCAGAAAGATTTTGTGCAACTGTTCATCATCAGTGTGGAGAGCTGATGCCAAGAA 793
QY      274 ArgProSerMetGlnGluIleValLysIleMet 284
      ::::::::::::::::::::
Db      794 CGGCATCATTTCAAGCAAAATCATTTCAATCTGT 826

```

```

RESULT 8
/ Sequence 4, Application US/09221236
/ Patent No. 6146841
/ GENERAL INFORMATION
/ APPLICANT: Acton, Susan
/ TITLE OF INVENTION: NOVEL CSApK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
/ FILE REFERENCE: NMI-050
/ CURRENT APPLICATION NUMBER: US/09/221,236
/ CURRENT FILING DATE: 1998-12-28
/ EARLIER APPLICATION NUMBER: 09/163,115
/ EARLIER FILING DATE: 1998-09-29
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
/ LENGTH: 2120
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (47)..(1411)
US-09-221-236-4

Alignment Scores:
Pred. No.: 1.02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 3 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-236-4 (1-2120)
QY 10 Serserserserseralaglygumctile-----glualaprosarginval 25
|||:::||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TCGACCCACCGCTCCGGTGGAGATATATACCTTTCATTATGAGATGCTCTCCGCT 61
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 26 leuansdhcgluglileasprfryllysgluileglualgluglualvalaiglyargly 45
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


Db	62	GCCTCCTTTGTGCAAAATAAATTGATGACTTCGAGCTTTTGTGAAACCTCGGCTGGAGGA	121
Qy	46	AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLysLys	63
Db	122	AGTTTGGGAGTGTTTATCGAGCAATGATATCACAGGCAAGGAGGTGGCTGTGAAG	181
Qy	64	-----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln	80
Db	182	AAGCTCTCAAAATAGAGAAGGCGAGAA-----ATA	214
Qy	81	LeuSerArgValAenHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro---	99
Db	215	CTCAGTCTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTGAACCTCC	274
Qy	100	---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly	118
Db	275	AACTATGGCATGTGCAGAAATATGCTTCTCTGGATCACTCTATGATTACATTAAACAGT	334
Qy	119	-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln	136
Db	335	AACAGAAGTAGGAGATG-----GATATGGATCACATTATGACCTGGGCCACTGAT	385
Qy	137	CysSerGlnGlyValAlaIleTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp	156
Db	386	GTAGCCAAAGAAATGCATTATTACATATGAGGCTCTCTCAAGGTGATTACAGAGAC	445
Qy	157	LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe	176
Db	446	CTCAAGTCAAGAAACGTTGTTATAGTCTGATGGA--GTACTGAAGATCTGTGACTTT	502
Qy	177	GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla	194
Db	503	GGT---GCCTCTCGGTTCCATAACACATACACACATGCTCTGGTTGGAATCTTCCCA	559
Qy	195	TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer	214
Db	560	TGGATGGCTCAGAAGTTATCCAGAGTCCCTCTGTGCAGAAACTTGTGCACATATCC	619
Qy	215	TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly	234
Db	620	TATGTGTGGTTCTCTGGAGATGCTAACAGGGAGGTCCCTTTTAAAGGTTTGGGAAGGA	679
Qy	235	ProAlaPheArgIleMetTrp--AlaValHisAsnGlyThrArgProProLeuIleLys	253
Db	680	-----TTACAAGTAGCTTGCTGTAGTGGAAAAAACGAGAGATTAAACATTCCACGC	733
Qy	254	AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln	273
Db	734	AGTTGCCCCAGAAGTTTGTCTGAACCTGTATACATCAGTGTGGGAAGCTGATGCCAAGAA	793
Qy	274	ArgProSerMetGluGluIleValLysIleMet	284
Db	794	CGGCATCATTTCAAGCAAAATCATTCATCTCGT	826

```

RESULT 9
US-09-221-416-4
; Sequence 4, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

RESULT 10
US-09-221-245-4
; Sequence 4, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-245-4

Alignment Scores:
Pred. No.: 1.02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 4 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-245-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCAGCGCTCGGTGAGAGTAAATTAATCTTGTGCATTATGAGATGTCCTGCTCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCCTTGTGCAATTAATTAATTGATGACTTGCACTTTTGAAGAACTGCGGTGAGAGA 121
QY 46 AlaPheGlyValAlaCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGCTTTGGAGGTGTTATTCAGACCAATGATATATCAAGACCAAGAGGTGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGGACAGA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGAGATATCTTGAACCTCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLysLysValAsnValLeuHisGly 118
DB 275 AACTATGGCATTGTACAGAAATATGCTCTCTGGGATACCTTATGATTAACATTAAACAGT 334
QY 119 -----AlaGluProLeuProTyrCThrAlaAlaHisAlaMetSerTyrCysLeuGln 136
DB 335 AACACAGTGAAGGAGATG-----GATATGATCTCATTAATGACCTGGGCCACAGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTAGCCAAAGGAATGCATTATTACATATGAGAGCTCCTGCAAGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLysLysIleCysAspPhe 176
DB 446 CTCAGGTCAAGAAAGCTGTGTTATAGCTCTGATGGA---GTACTGAAGATCTGTGACTCTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnLeuLys-----GlySerAlaAla 194
DB 503 GGT---GCCTCGGTCCATACCATATCAACACACATGTCCTTGGTTGGAACCTTCCCA 559

```

```

QY 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGGATGGCTCCAGAAAGTTATTCAGAGTCCTCGTGTACAGAACTTGTGACATTTCC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyGly 234
DB 620 TATGTTGTGTTCTCTGGGAGATGCTAACAGGAGGTCCCTTTAAAGTTTGAAGGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProLeuIleLys 253
DB 680 -----TTACAAGTAGCTTGCTGTAGTGGAAGAAAAACAGAGATTAACTTCCAAAC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysThrSerLysAspProSerGln 273
DB 734 AGTTGCCCAAGAGTTTGTGTAACGTTAACATCATGTTGGAGACTGATGCCAAGAAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMet 284
DB 794 CGGCATCATTCAGCAAAATCATTCATTCCTG 826

RESULT 11
US-09-163-115-4
; Sequence 4, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-163-115-4

Alignment Scores:
Pred. No.: 1.02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 4 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-163-115-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCGACCCAGCGCTCGGTGAGAGTAAATTAATCTTGTGCATTATGAGATGTCCTGCTCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCCTTGTGCAATTAATTAATTGATGACTTGCACTTTTGAAGAACTGCGGTGAGAGA 121
QY 46 AlaPheGlyValAlaCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGGTGTTATTCAGACCAATGATATATCAAGACCAAGAGGTGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGGACAGA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGAGATATCTTGAACCTCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLysLysValAsnValLeuHisGly 118

```

Db 275 AACTATGACATGTCACAGAAATATGCTTCTCTGGATCAGCTCTATGATACATTAACAGT 334
Qy 119 -----AlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAACTGAGGAGATG-----GATATGGATCACATTAATGACCTGGGCCACTGAT 385
Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCAATTAATACATATGAGGCTCTGTCACAGGATGATTCACAGAGAC 445
Qy 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAACTGTTATAGCTCTGATGGA-----GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATAACCATACACACACATGCTCTTGGTTGGAATTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAACTGATCCAGAGTCTCCCTGTGTGAGAACTTTAAAGGTTTGAAGGA 679
Qy 215 TrpGlyIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyGly 234
Db 620 TATGGTGGTCTCTCTGGAGATGCTAACAGGAGGTCCTCTTAAAGGTTTGAAGGA 679
Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProLeuLys 253
Db 680 -----TTCAAGTAGCTTGGCTTGTAGTGGAAAAAACAGAGATTAACCATTTCCAAAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTGTGCCAGAACTTTCTGCAACTGTTTACATCATCTAGTGTGGAAAGCTGATGCCAAGAA 793
RESULT 12
US-09-221-528-4
; Sequence 4, Application US/09221528
; Patent No. 6150874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-528-4
Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 4 Gaps: 12
US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-528-4 (1-2120)
Qy 10 SerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25

Db 2 TCGACCCACGCGTCCGGTGGAAAGTATAATCTTTGTCATTATGAGATGTCGCTCTCGGT 61
Qy 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValValGlyArgGly 45
Db 62 GCCTCTCTTGTGCAAAATTAATTTGATGACTTGCAGCTTTTGAAGAACTGCGGTGGAGA 121
Qy 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaLys 63
Db 122 AGTTTGGAGTGTTCATCGAGCCAAATGATATCACAGGACAAAGAGGTGCTGTAAAG 181
Qy 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
Db 182 AAGCTCTCTCAAAATAGAGAAAGAGGAGAA-----ATA 214
Qy 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
Db 215 CTCAGTGTCTCTAGTCACAGAAACATCATCCAGTTTATGGAGTAATCTTGAACCTCCC 274
Qy 100 ---ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGly 118
Db 275 AACTATGGCATTGTACAGAAATATGCTTCTCTGGATCACTCTATGATTACATTACAGT 334
Qy 119 -----AlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
Db 335 AACAGAACTGAGGAGATG-----GATATGGATCACATTAATGACCTGGGCCACTGAT 385
Qy 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
Db 386 GTAGCCAAAGGAATGCAATTAATTTACATATGAGGCTCTCTGTCAAGGTGATTCACAGAC 445
Qy 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
Db 446 CTCAAAGTCAAGAACTGTTTATAGCTGTCTGATGGA-----GTACTGAAGATCTGTGACTTT 502
Qy 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
Db 503 GGT---GCCTCTCGGTTCCATAACCATACACACACATGCTCTTGGTTGGAATTTCCCA 559
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 560 TGGATGGCTCCAGAACTTATCCAGAGTCTCCCTGTGTGAGAACTTTGTACACATATTC 619
Qy 215 TrpGlyIleLeuTrpGluValIleThrArgLysProPheAspGluIleGlyGly 234
Db 620 TATGGTGGTCTCTCTGGAGATGCTAACAGGAGGTCCTCTTAAAGGTTTGAAGGA 679
Qy 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProLeuLys 253
Db 680 -----TTCAAGTAGCTTGGCTTGTAGTGGAAAAAACAGAGATTAACCATTTCCAAAGC 733
Qy 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
Db 734 AGTGTGCCAGAACTTTGTGCAACTGTTTACATCATCTAGTGTGGAAAGCTGATGCCAAGAA 793
Qy 274 ArgProSerMetGluGluIleValLysIleMet 284
Db 794 CGGCCATCATTCAGCAAAATCATTTCAATCTCTG 826
RESULT 13
US-09-593-553-4
; Sequence 4, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0

```

; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-593-553-4

Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-593-553-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCACCCACGCGTCGCGGAGATGTAATTAATTGTCATTATGAGATGTCGTCCTCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCTTTGTCACAAATTAATTTGATGACTTTTGGAAAACTCGCGGTGAGGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGCTTTTATCGAGCCAAATGATATCATCAGCAAGAGAGAGGTGCTGAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGCGAGAA-----ATA 214
QY 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTTCAGTCACAGAAACATCATCCAGTTTATGAGAAATCTCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
DB 275 AACTATGGCATTTGTCAGAAATATGCTTCTGTGGATCATCTTATGATTACATTAAACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
DB 335 AACGAAAGTGAAGAGATG-----GATATGATCATCATTAATGACCTGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTACCCAAAGGAATGCAATTATTACATATGAGGCTCCTGTCAAGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
DB 446 CTCAGATCAAGAAACGTTTATAGCTGTGATGA---GTACTGAAGATCTGTGACTTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
DB 503 GGT---GCCTCTGGTTCATTAACCATACACACATGTCCTTGGTGAACCTTCCCA 559
QY 195 TyrMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGGATGGCTCCAGAAATTATCCAGAGTCTCCCTGTCAAGAACTTGACACATATTC 619
QY 215 TyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 234
DB 620 TATGTGTGCTTCTCTGAGATGCTTAACAAAGAGAGTCCCTTTTAAAGGTGGAGGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAAAGTAGCTGTGATGAGAAAAAAGAGAGATTAAACCATTCACAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
DB 734 AGTTGCCCAAGATTGCTGAAGTTCATCATCAGTGTGGAGAGCTATGCCAAGAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMet 284
DB 794 CGGCATCATTCAGCAAAATCATTTCAATCTCG 826

```

```

RESULT 14
US-09-221-237-4
; Sequence 4, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-237-4

Alignment Scores:
Pred. No.: 1,02e-43 Length: 2120
Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-221-237-4 (1-2120)
QY 10 SerSerSerSerSerAlaGlyGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCACCCACGCGTCGCGGAGATGTAATTAATTGTCATTATGAGATGTCGTCCTCGGT 61
QY 26 LeuAsnPheGluGluIleAspTyrLysGluIleGluValGluGluValAlaGlyArgGly 45
DB 62 GCCTCTTTGTCACAAATTAATTTGATGACTTTTGGAAAACTCGCGGTGAGGA 121
QY 46 AlaPheGlyValValCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGAGCTTTTATCGAGCCAAATGATATCATCAGCAAGAGAGGTGCTGAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPheIleValGluLeuArgGln 80
DB 182 AAGCTCCTCAAAATAGAGAAAGAGCGAGAA-----ATA 214
QY 81 LeuSerArgValAlaAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro--- 99
DB 215 CTCAGTGTCTTCAGTCACAGAAACATCATCCAGTTTATGAGAAATCTCTTGAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
DB 275 AACTATGGCATTTGTCAGAAATATGCTTCTGTGGATCATCTTATGATTACATTAAACAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
DB 335 AACGAAAGTGAAGAGATG-----GATATGATCATCATTAATGACCTGGCCACTGAT 385
QY 137 CysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTACCCAAAGGAATGCAATTATTACATATGAGGCTCCTGTCAAGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
DB 446 CTCAGATCAAGAAACGTTTATAGCTGTGATGA---GTACTGAAGATCTGTGACTTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
DB 503 GGT---GCCTCTGGTTCATTAACCATACACACATGTCCTTGGTGAACCTTCCCA 559
QY 195 TyrMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGGATGGCTCCAGAAATTATCCAGAGTCTCCCTGTCAAGAACTTGACACATATTC 619
QY 215 TyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGly 234
DB 620 TATGTGTGCTTCTCTGAGATGCTTAACAAAGAGAGTCCCTTTTAAAGGTGGAGGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAAAGTAGCTGTGATGAGAAAAAAGAGAGATTAAACCATTCACAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
DB 734 AGTTGCCCAAGATTGCTGAAGTTCATCATCAGTGTGGAGAGCTATGCCAAGAA 793
QY 274 ArgProSerMetGluGluIleValLysIleMet 284
DB 794 CGGCATCATTCAGCAAAATCATTTCAATCTCG 826

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 02:09:35 ; Search time 98 Seconds
(without alignments)
1205.608 Million cell updates/sec

Title: US-09-830-144-2_COPY_1_303

Perfect score: 1615

Sequence: 1 MSTAASASSSSSSAGEMIE.....MTHLMRYFFGADEPLQVPCQ 303

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DBV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002.141512.2975/app.query.fasta_1.455
-DB=Published Applications NA -QWMT=fastap -SUFFX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNING=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144 @CNG 1.1.36 @runat_04122002.141512.2975
-NCPU=6 -ICPU=3 -NO_XLPX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1615	100.0	1788	9	US-10-158-895-14
2	1615	100.0	2656	9	US-10-158-895-3
3	482.5	29.9	3454	10	US-09-969-347-226
4	442	27.4	2120	10	US-09-757-982-4

5	433	26.8	1365	10	US-09-757-982-6	Sequence 6, Appli
6	432	26.7	3111	12	US-10-014-882-1	Sequence 1, Appli
7	432	26.7	3518	12	US-10-014-882-3	Sequence 3, Appli
8	411.5	25.5	1662	9	US-09-938-842A-1014	Sequence 1014, Ap
9	410	25.4	2505	10	US-09-947-199-3	Sequence 3, Appli
10	410	25.4	2505	10	US-09-947-199-9	Sequence 9, Appli
11	410	25.4	3025	10	US-09-947-199-1	Sequence 1, Appli
12	410	25.4	3026	10	US-09-947-199-7	Sequence 7, Appli
13	392	24.3	3286	10	US-09-904-389-1	Sequence 1, Appli
14	391.5	24.2	2211	9	US-09-938-842A-1577	Sequence 1577, Ap
15	387	24.0	1428	9	US-09-938-842A-882	Sequence 882, App
16	387	24.0	2892	9	US-09-938-842A-1073	Sequence 1073, Ap
17	369	22.8	1627	10	US-09-828-313-16	Sequence 16, Appl
18	365.5	22.6	3663	10	US-09-919-172-84	Sequence 84, Appl
19	358.5	22.2	6383	9	US-09-954-531-405	Sequence 405, App
20	357.5	22.1	2574	10	US-09-735-103-2	Sequence 2, Appli
21	357.5	22.1	2574	12	US-10-045-428A-2	Sequence 2, Appli
22	351.5	21.8	1638	9	US-09-938-842A-903	Sequence 903, App
23	345.5	21.4	2661	9	US-09-938-842A-1085	Sequence 1085, Ap
24	340.5	21.1	2719	10	US-09-917-800A-1405	Sequence 1405, Ap
25	333	20.6	2000	10	US-09-977-269-1	Sequence 1, Appli
26	330	20.4	2977	10	US-09-969-347-207	Sequence 207, App
27	329.5	20.4	4165	8	US-08-578-684-1	Sequence 1, Appli
28	329	20.4	2456	9	US-10-186-399-1	Sequence 1, Appli
29	329	20.4	2500	10	US-09-977-269-3	Sequence 3, Appli
30	329	20.4	3454	12	US-10-044-090-48	Sequence 48, Appl
31	329	20.4	3726	10	US-09-925-302-271	Sequence 271, App
32	328.5	20.3	1092	9	US-09-938-842A-553	Sequence 553, App
33	327.5	20.3	3012	9	US-09-938-842A-1479	Sequence 1479, Ap
34	321.5	19.9	1215	10	US-09-880-107-2254	Sequence 2254, Ap
35	320	19.8	3791	10	US-09-757-100B-1	Sequence 1, Appli
36	319.5	19.8	3945	10	US-09-921-771-4	Sequence 4, Appli
37	319.5	19.8	3969	10	US-09-982-610-23	Sequence 23, Appl
38	318.5	19.7	1911	10	US-09-917-800A-1611	Sequence 1611, Ap
39	318.5	19.7	3107	10	US-09-954-456-1128	Sequence 1128, Ap
40	318.5	19.7	3348	10	US-09-982-610-34	Sequence 34, Appl
41	313.5	19.4	4698	10	US-09-895-652-1	Sequence 1, Appli
42	312.5	19.3	3393	10	US-09-922-138-3	Sequence 3, Appli
43	312.5	19.3	3633	10	US-09-922-138-1	Sequence 1, Appli
44	311.5	19.3	2015	10	US-09-954-456-1983	Sequence 1983, Ap
45	308.5	19.1	2442	10	US-09-827-949-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-10-158-895-14
; Sequence 14, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR FILING DATE: 2000-04-11
; PRIOR FILING DATE: 1998-10-22
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-10-158-895-14

Alignment Scores:

```

Pred. No.: 7.64e-1
Score: 1615.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9

```

US-09-830-144-2_COPY 1_303 (1-303) X US-10-158-895-14 (1-1788)

[illegible]

Db 907 CCTTGTGTCAG 915

RESULT 2

US-10-158-895-3
 / Sequence 3, Application US/10158895
 / Patent No. US20020155624A1
 / GENERAL INFORMATION:
 / APPLICANT: ONO, KOICHIRO
 / APPLICANT: OHTOMO, TOSHIIKO
 / APPLICANT: TSUCHIYA, MASAYUKI
 / TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
 / FILE REFERENCE: 053466/0278
 / CURRENT APPLICATION NUMBER: US/10/158,895
 / CURRENT FILING DATE: 2002-06-03
 / PRIOR APPLICATION NUMBER: US/09/529,279
 / PRIOR FILING DATE: 2000-04-11
 / PRIOR APPLICATION NUMBER: PCT/J998/04796
 / PRIOR FILING DATE: 1998-10-22
 / PRIOR APPLICATION NUMBER: JP 9/290188
 / PRIOR FILING DATE: 1997-10-22
 / NUMBER OF SEQ ID NOS: 48
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 3
 / LENGTH: 2656
 / TYPE: DNA
 / ORGANISM: Homo sapiens
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: (183) ..(1919)
 / US-10-158-895-3

Alignment Scores:

```
Pred. No.: 1.39e-1
Score: 1615.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 9
```

US-09-830-144-2_COPY_1_303 (1-303) x US-10-158-895-3 (1-2656)

```

QY      1 MetSerThrAlaSerAlaAlaSerSerSerSerSerSerAlaGlyGlyMetLeuLeu 20
Db      183 AAGTTAACAGCCTCGCGGCTCTCTCTCTCTCGTTCGGCGGGTGAATGATCGAA 242
QY      21 AlaProSerGlnValLeuAsnPhenGlyGlnLeuAspTyrLeuGlyLeuValGlyGln 40
Db      243 GCCCTTCCAGGCTCTCAACTTTGGAAGAGTGCATCAAGAGAGATCGAGCTGGAAGAG 302
QY      41 ValValGlyArgGlyAlaPheGlyValValCysLeuAlaLysTrpArgAlaLysAspVal 60
Db      303 GTTGTGGGAAGAGAGCCTTTGGAGATTGTTGGCAACCTAAGTGGAGAGCAAAAGATGTT 362
QY      61 AlaIleLysGlnIleGluSerGlnSerGlnArgLysAlaPheIleValGluLeuArgGln 80
Db      363 GCATTTAAACCAATGAAAGAGATCTCGAGAGAAAGCCTTTATTGTAGAGCTTCCGAG 422
QY      81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnProVal 100
Db      423 TTTTCCCGGTGAACCATCTTAATTTGTAAAGCTTTATGAGCGCTCTGGAATCCAGTG 482
QY      101 CysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAlaGlu 120
Db      483 TGTCTTGTATGAATATAGCTGAAGGGGGCTTTATATATATGTCGCATGGTCTGAA 542
QY      121 ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnGlyCysSerGlnGly 140
Db      543 CCAATCCCATATTAATACGTGCTCCCAACCCATGAGTGTTGTATTACAGTGTTCCCAAGGA 602
QY      141 ValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProPro 160
Db      603 GTGGCTTATCTTCACAGCATGCACCCAAAGCGCTAATTCCAGAGGAGCTGMAACCA 662

```



```
Qy 161 AsnLeuLeuValAlaGlyThrValLeuLysIleCysAspPheGlyThrAlaCys 180
Db 663 AACTTACTGCTGCTGAGGGGGGACAGTTCTAAATAATTGTGATTTGGTACAGCCTGT 722
Qy 181 AspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluVal 200
Db 723 GACATTACAGACACATGACCAATAACAAGGGAGTGCTTGGATGGACCTGAAGTT 782
Qy 201 PheGluGlySerAsnTrpSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrp 220
Db 783 TTTGAAGGTAGTAATTACAGTGAATAATGTGAGCTCTTCAGCTGGGGTATTATTCTTTGG 842
Qy 221 GluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMet 240
Db 843 GAAAGTATACCGCTCGGAAACCTTTGATGAGATTGGTGGCCAGCTTTCCGAATCATG 902
Qy 241 TrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGlu 260
Db 903 TGGGCTGTTCAATAATGGTACTCCACCACCACTGATAAAAAATTACCTAAGCCCATTTGAG 962
Qy 261 SerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIle 280
Db 963 AGCTGATGACTCGTTGGTCTAAGATCCTTCCAGCGCCCTTCAATGAGGAATT 1022
Qy 281 ValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyr 300
Db 1023 GTGAAAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGATGAGCCATTACAGTAT 1082
Qy 301 ProCysGln 303
Db 1083 CCTGTGCAG 1091
```

RESULT 3

```
US-09-969-347-226
; Sequence 226, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 226
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-226
```

Alignment Scores:

```
Pred. No.: 8,03e-47 Length: 3454
Score: 482.50 Matches: 107
Percent Similarity: 56.63% Conservative: 51
Best Local Similarity: 38.35% Mismatches: 98
Query Match: 29.88% Indels: 23
DB: 10 Gaps: 7
```

US-09-830-144-2_COPY_1_303 (1-303) x US-09-969-347-226 (1-3454)

```
Qy 21 AlaProSerGlnValLeuAsnPheGluGluIleAspTrpLysGluIleGluValGluGlu 40
Db 535 GCACCGCGGGCCCTCCAGCTGCCCGCAGGATGCCCTTCCACGAGCTGCAGCTAGAGGAG 594
Qy 41 ValValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspVal 60
Db 595 ATCATCGTGTGGGGGGCTTTTGGCAAGTCTATCGGCGCCCTGTGGCGTGGCGAGGAGTG 654
```

```
Qy 61 AlaIleLysGlnIleGluSerGluSerGluArgLysAlaPheIleVal----- 76
Db 655 GCAGCTAAGCGCGCGCGCTGGAGCCCTTGAGAGAGACCCGCGCAGTGACACGCGAGCAGGTG 714
Qy 77 -----GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly 94
Db 715 TGCAGAGAACCCCGGCTCTTTGGAGCCCTGCAGCACCCCAACATAATTGCGCTTAGGGGC 774
Qy 95 AlaCysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeu 112
Db 775 GCCTGCTCAACCCCGCCACACCTCTGCTAGTGATGAGTATGCCCGGGGTGGTGCACTG 834
Qy 113 TyrAsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 132
Db 835 AGCAGGGTGTCTGGCAGGTGCGCGGGTGCACCTCCACCTG-----CTGGTCAAC 882
Qy 133 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 152
Db 883 TGGCTGTGCAAGGTGCGCGGGCATGAACCTACCTACAAATGATGCCCTGTGCCCATC 942
Qy 153 IleHisArgAspLeuLysProProAsnLeuLeuVal----- 165
Db 943 ATCCACCGGACCTCAAGTCCATCAACATCCTGATCCTGGAGGCCATCGAGAACCAAC 1002
Qy 166 AlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThr 184
Db 1003 CTCGCAGACACGGTGCTCAAGATCACGGAATTCGCGCCTCGCCGCGAGTGCGCACAGACC 1082
Qy 185 HisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySer 204
Db 1063 ACCAAGATGAGCGCTGCGGGACCTACGCTGGATGGCGGAGGTTATCCGCTCTCTCC 1122
Qy 205 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThr 224
Db 1123 CTCCTTCACAAAGCAGTGAAGTCTGGAGCTTCGGGGTGTGCTGTGGAGCTCTGACG 1182
Qy 225 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaVal--- 243
Db 1183 GGGGAGGTCCCTACCGGTGAGATC-----GACGCTTTGGCGGTGCGGTATGGCGTGCT 1236
Qy 244 HisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMet 263
Db 1237 ATGAATAAGCTGACGCTGCCCATTCCTCCACGTGCCCGAGCCCTTTGCCCGCTCTCTG 1296
Qy 264 ThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 282
Db 1297 GAGGATGCTGGGACCCAGACCCCGGGCGGCAGATTTCGGTAGCATCTTGAAG 1353
RESULT 4
US-09-757-982-4
; Sequence 4, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-757-982-4
Alignment Scores:
Pred. No.: 2.3e-42 Length: 2120
```

Score: 442.00 Matches: 105
Percent Similarity: 55.33% Conservative: 56
Best Local Similarity: 36.08% Mismatches: 98
Query Match: 27.37% Indels: 32
DB: 10 Gaps: 12

US-09-830-144-2_COPY_1_303 (1-303) x US-09-757-982-4 (1-2120)

QY 10 SerSerSerSerSerAlaGluGluMetIle-----GluAlaProSerGlnVal 25
DB 2 TCAGCCACGCGTCGCGTGAAGTAAATCTTGTTCATTATGAGATGCGCTCTCGGT 61
QY 26 LeuAsnPhcGluGluIleAspTyrLysGluIleGluValGluValAlaGlyArgGly 45
DB 62 GCCTCCTTTGTGCAAAATTAAATTGATGACTTCGACGTTTGGAAACTGCGGTGAGGA 121
QY 46 AlaPhcGluValAlaCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys 63
DB 122 AGTTTGGGAGTGTATTATGAGCCAAATGATATCAACAGACAGAGAGGTGCTGTAAAG 181
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPhcIleValGluLeuArgGln 80
DB 182 AAGCTCTCCAAATAGAGAAAGGAGCAGAA-----ATA 214
QY 81 LeuSerArgValAsnHisProAsnIleValLysLeuTyrGluValAlaCysLeuAsnPro--- 99
DB 215 CTCAGTCTCTCTGACAGAAACATCATCCAGTTTATGAGTAATTCCTTAACCTCCC 274
QY 100 ---ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly 118
DB 275 AACTATGAGATGTCACAGAAATATGCTTCTCGGATCATCTTAAGATATACATTAAAGT 334
QY 119 -----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGln 136
DB 335 AACAGAAAGAGAGAGATG-----GATATGATCATCATTAATTAACCTGGGCCACTGAT 385
QY 137 CysSerGluGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAsp 156
DB 386 GTAGCCAAAGAAATGATTAATTACATATGAGAGCTCCTGTCAAGGTGATTCACAGAGAC 445
QY 157 LeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhe 176
DB 446 CTCCAATCAAGAAACGTTGTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTT 502
QY 177 GlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAla 194
DB 503 GGT---GCCTCTCGGTTCATACATACACACACATGCTCTGTGTAACCTTCCCA 559
QY 195 TrpMetAlaProGluValAlaPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
DB 560 TGGATGGCTCCAGAAATTATCCAGAGCTCCTGTGTCAGAAACTTGTGACACATATTCC 619
QY 215 TrpGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGly 234
DB 620 TATGGTGTGTTCTCTCGGAGATGCTAACAGAGGAGGTCCCTTTAAAGGTTTGGAAAGA 679
QY 235 ProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLys 253
DB 680 -----TTACAAGTAGCTGGCTTGTAGTGAAGAAAAAGAGAGATTAAACCATTCACAGC 733
QY 254 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 273
DB 734 AGTTGCCCCAGAGATTGTGCTGAACGTGTATCATCAGTGTGGAGAACTGTATGCCAAGAA 793
QY 274 ArgProSerMetGluGluIleValValLysIleMet 284
DB 794 CGGCCATCATTTCAAGCAATTCATTCAATCTCG 826

APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSARK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: MWI-050
CURRENT APPLICATION NUMBER: US/09/757,982
CURRENT FILING DATE: 2001-01-10
PRIORITY APPLICATION NUMBER: 09/163,115
PRIORITY FILING DATE: 1998-09-29
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 1365
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1365)
US-09-757-982-6

Alignment Scores:
Pred. No.: 1,366-41 Length: 1365
Score: 433.00 Matches: 99
Percent Similarity: 56.30% Conservative: 53
Best Local Similarity: 36.67% Mismatches: 90
Query Match: 26.81% Indels: 28
DB: 10 Gaps: 11

US-09-830-144-2_COPY_1_303 (1-303) x US-09-757-982-6 (1-1365)

QY 27 AsnPhcGluGluIleAspTyrLysGluIleGluValGluValAlaGlyArgGlyAla 46
DB 19 TCCTTTGTGAATTAATTGATGACTTCGACTTTTGGAAACGCGGTGAGAGAACT 78
QY 47 PhcGlyValAlaCysLysAlaLysTrp-----ArgAlaLysAspValAlaIleLys--- 63
DB 79 TTTGGAGTGTATTATGAGCCAAATGATGATATCAACAGACAGAGAGGTGCTGTAAAGAG 138
QY 64 -----GlnIleGluSerGluSerGluArgLysAlaPhcIleValGluLeuArgGlnLeu 81
DB 139 CTCCTCAAAATAGAGAAAGGAGCAGAA-----ATATCTC 171
QY 82 SerArgValAsnHisProAsnIleValLysLeuTyrGluValAlaCysLeuAsnPro----- 99
DB 172 AGTGTCTCAGTACAGAAACATCATCCAGTTTATGAGATTAATTTGAACCTCCCAAC 231
QY 100 ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnValLeuHisGly--- 118
DB 232 TATGGCATGTGTACAGAAATGCTTCTCGGATCACTCATGATATTAACATTAACAGTAA 291
QY 119 ---AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
DB 292 AGAAGTAGAGAGATG-----GATATGATCATCATTAATTAACCTGGGCCACTGATGTA 342
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
DB 343 GCCAAAGAAATGATTAATTACATATGAGAGCTCCTGTCAAGATGATTAACAGAGCTTC 402
QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPhcGly 177
DB 403 AAGTCAAGAAACGTTGTATAGCTGCTGATGGA---GTACTGAAGATCTGTGACTTTGGT 459
QY 178 ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----GlySerAlaAlaTrp 195
DB 460 ---GCCCTCGGTTCCATTAACATTAACACACATGCTCTGTGTAACCTTCCCATGG 516
QY 196 MetAlaProGluValAlaPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215
DB 517 ATGGCTCCAGAAATTATCCAGAGCTCCTGTGTCAAGAACTTGTGACACATATTCTAT 576
QY 216 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
DB 577 GGTGTGTCTCTCGGAGATGCTTAACAGAGAGGTCCCTTTAAAGGTTTGGAAAGA--- 633
QY 236 AlaPhcArgIleMetTrp---AlaValHisAsnGlyThrArgProProLeuIleLysAsn 254

```
Db 634 ---TTACAAGTAGCTTGGCTTTGTAGTGGAAGAAAAACGAGAGATTAAACCAATTCCAAGCAGT 690
Qy 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
Db 691 TGCCCCAGAAGTTTGTGAAGCTGTATACATGCTGTGGAGAGCTGTGGCAAGCTGATGCAAGAACGG 750
Qy 275 ProSerMetGluGluIleValLysIleMet 284
Db 751 CCATCATTTCAAGCAATCATTTCAATCTCTG 780

RESULT 6
US-10-014-882-1
; Sequence 1, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-014-882-1

Alignment Scores:
Pred. No.: 6,21e-41 Length: 3111
Score: 432.00 Matches: 99
Percent Similarity: 53.61% Conservative: 57
Best Local Similarity: 34.02% Mismatches: 107
Query Match: 26.75% Indels: 28
DB: Gaps: 7

US-09-830-144-2_COPY_1_303 (1-303) x US-10-014-882-1 (1-3111)

Qy 22 ProSerGlnValLeuAsnPhleGluIleAspTyrLysGluIleValGluVal 41
Db 328 CCCCTCGCGGCCAGCTCCCGGTTACAGTCCGCTTCGAGCGGCTGGAGCTGAAGAGCTC 387
Qy 42 ValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspValAla 61
Db 388 ATCGCGCGTGGGGCTTCGGGACAGTGTACCGGCCACCTGGCAGGCGCCAGGAGGTGGCC 447
Qy 62 IleLysGlnIleGluSerGluArgLysAlaPheIleVal----- 76
Db 448 GTGAAGCGCGCCAGGACCCGAGACGAGACGCGCGCGCTGCCGAGCGGTGGCG 507
Qy 77 ---GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 95
Db 508 CGCGAGGCTCGGCTTCGCCATGCTCGGCACCCCAACATCATCGAGCTGCGCGCGGTG 567
Qy 96 CysLeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 113
Db 568 TGCCTGACGACGCGCACCTTCGCTGCTGTGAGTTCGCGCGCGCGCGGCGCTCAAC 627
Qy 114 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAla----- 127
Db 628 CGAGCGTGGCGGCTGCCAAGCGCGCCCGGACCCGCGCGCGCGCGCGCGCGCG 687
Qy 128 -----AlaHisAlaMet---SerTrpCysLeuGlnCysSerGlnGlyValAla 142
Db 688 CGCGCATCCCTCCGACAGTGTGTCACTGGGCCGTGCAGATAGCGCGGGCATGCTC 747
Qy 143 TyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProAsnLeu 162
Db 143 TyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysProAsnLeu 162
```

```
Db 748 TACCTGCATGAGGAGCGCTTCGTGCCCATCTCTGCACCGGACCTCAAGTCCAGCAACATT 807
Qy 163 LeuLeuValAlaGly-----GlyThrValLeuLysIleCysAsp 175
Db 808 TTGCTACTTTGAGAAGATGAACATGATGACATCTGCAATAAAACTTTGAAGATTACAGAT 867
Qy 176 PheGlyThrAlaCysAsp---IleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 194
Db 868 TTTGGGTTGGCGAGGAATGGCACAGGACCCAAATAATGACACAGCAGGACCTATGCC 927
Qy 195 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 214
Db 928 TGGATGGCCCCGGAAGTATCAAGTCTTCTTGTCTTCTAAGGGAAGCGACATCTGGAGC 987
Qy 215 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 234
Db 988 TATGAGTGCTCTGTGGAACTGCTCACCGGAGAGTCCCTCATCGGGGATTTGATGCG 1047
Qy 235 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 254
Db 1048 CTCGCGGTGGCTTATGGGTAGCAGTCAATAACTCACTTTG---CCCATTCATCCACC 1104
Qy 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
Db 1105 TGCCTCGAGCGCTTTGCCAAGCTCATGAAGAAATGCTGGCAACAGACCTCATATTG 1164
Qy 275 ProSerMetGluGluIleValLysIleMetThr 285
Db 1165 CCATCGTTGCTTAATTCGAACAGTTGACT 1197

RESULT 7
US-10-014-882-3
; Sequence 3, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3518
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-014-882-3

Alignment Scores:
Pred. No.: 7,48e-41 Length: 3518
Score: 432.00 Matches: 99
Percent Similarity: 53.61% Conservative: 57
Best Local Similarity: 34.02% Mismatches: 107
Query Match: 26.75% Indels: 28
DB: Gaps: 7

US-09-830-144-2_COPY_1_303 (1-303) x US-10-014-882-3 (1-3518)

Qy 22 ProSerGlnValLeuAsnPhleGluIleAspTyrLysGluIleValGluVal 41
Db 591 CCCCTCGCGGCCAGCTCCCGGTTACAGTCCGCTTCGAGCGGCTGGAGCTGAAGGAGCTC 650
Qy 42 ValGlyArgGlyAlaPheGlyValValCysLysAlaLysTrpArgAlaLysAspValAla 61
Db 651 ATCGCGCGTGGGGCTTCGGGACAGTGTACCGGCCACCTGGCAGGCGCCAGGAGGTGGCC 710
Qy 62 IleLysGlnIleGluSerGluArgLysAlaPheIleVal----- 76
Db 711 GTGAAGCGCGCGCGGACCCGAGACGCGCGCGCGGCTGCCGAGCGGTGGCG 770
```



```

RESULT 9
US-09-947-199-3
; Sequence 3, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
US-09-947-199-3
Alignment Scores:
Pred. No.: 1.76e-38 Length: 2505
Score: 410.00 Matches: 103
Percent Similarity: 57.89% Conservative: 51
Best Local Similarity: 38.72% Mismatches: 91
Query Match: 25.39% Indels: 22
DB: 9 Gaps: 9

US-09-830-144-2_COPY_1_303 (1-303) x US-09-947-199-3 (1-2505)
Qy 31 IleAspTyrLysGluLeuValGluValAlaPheGlyAlaPheGlyValVal 50
Db 1372 CTTCAGCTCTCAGAAATTGAGTTCCATGAGATTATTGGCTCAGGTCTTTTGGGAAGTA 1431
Qy 51 CysIysAlaLysTrpArgAlaLysAspValAlaLeuLysGlnLeuGlu----- 66
Db 1432 TATAAAGGACGATGCAGAAATAAAAATAGTGCTATAAAACTTATCGAGCCAATACCTAC 1491
Qy 67 ---SerGluSerGluArgLysAlaPheLeuValGluLeuArgGlnLeuSerArgValAsn 85
Db 1492 TGCTCCCAAGTCAGATGGGATGATGTTGGCCGAGAGGTGCCATTCTCGCCACTCAAT 1551
Qy 86 HisProAsnIleValLysLeuTyrGlyAlaCysLeuAsn-----ProValCysLeu 102
Db 1552 CATCCCTCGTAATTCAGTTTGTGGTGCTTGTGAATGATCCAGCCAGTTTGGCCATT 1611
Qy 103 ValMetGluTyrAlaGluGlyCysLeuTyrAsnValLeuHisGlyAlaGluProLeu 122
Db 1612 GTCACTCAATCATATCAGGGGGTCTCTGTTCTCCCTCTCATGA-GCAGAAGAGGAT 1670
Qy 123 ProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln-GlyValAl 142
Db 1671 TCCTTGATT-----GCAGTCTAAATTAAATTATTTGCAGTAGATGTTGCCAAGGCATGGA 1724
Qy 142 aTyrLeuHisSerMet---GlnProLyseAlaLeuIleHisArgAspLeuLysProAs 161
Db 1725 GTACCTTCACAACTTGACACAGCCA-----ATTATATCATGTCGACTTTGAACATCACAA 1778
Qy 161 nLeuLeuLeuValAlaGlyGlyThrValLeuLysLeuLysCysAspPheGlyThrAlaCysAs 181
Db 1779 TATTCTTCTATGAGGATGGCATGCTGTG--GTGCGCATTTTGGAGAAATCAAGATT 1835
Qy 181 piledInThr-----HisMetThraAsnLysGlySerAlaLaITrpMetAl 197
Db 1836 TCTACAGTCTCTGGATGAAGACAACATACAAAAACAACCTGGGAACCTCCGTTGCTGATGGC 1895

```

```
Db 1519 TGCCGAGAGGTGTCCATTCCTCTGCCAGCTCAACACCCTCGGTGGTTCAGTTTGTGGGT 1578
Qy 95 AlaCysLeu-----AsnProValCysLeuValMetGluTyrAlaGluGlySer 111
Db 1579 GCGTCCTGGATGACCCCGATGCTAGTTGGCATTCACACAGTTCCTCAGAGAGGCTCC 1638
Qy 112 LeuTyrAsnValLeuHISGlyAlaGluProLeuProTyrTyrThrAlaHisAlaMet 131
Db 1639 CTGTTCTCCCTGCTTCAT-----GAACAGAGAGAAATTCCTGACTTGCAGTCTTAATTA 1692
Qy 132 SerTyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMet--GlnProlys 150
Db 1693 ATCATTTGCGGTAGACGTTCGCAAGGCGATGAGTACTCTCACAGCTTGAACCCAGCA--- 1749
Qy 151 AlaLeuHISHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrVal 170
Db 1750 ---ATCATACACCGCAGCAGCTGAACAGCCCAATATTCCTCTCATAGAGATGGCATGCT 1806
Qy 171 LeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMet 186
Db 1807 GTG---GTGGCAGATTTTGGAGATCAAGATTTTCTCAGTCCCTGGATGAGATGAACATG 1863
Qy 187 ThrAsnAsnLysGlySerAlaAlaTyrMetAlaProGluValPhe--GluGlySerAsn 205
Db 1864 ACAAGACAGCCGAGGAACCTCGCGCTGATGGCCCCCTGAGGTTCACACAGTGCAGAGA 1923
Qy 206 TyrSerGluLysCysAspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArg 225
Db 1924 TACACCATCAAGCGTATGCTTCAGTTACTCCTGTGTCTGTGGAGAGCTCTCACTGGA 1983
Qy 226 ArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTyrAlaHisAsn 245
Db 1984 GAAATTCATTCCTCATCTCAACAGCCGCGTGCAGCAGCAGATGTGCGATACCAACC 2043
Qy 246 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 265
Db 2044 ---ATCAGACCGGCCCATCGGCTATTCATCCCAAGCCCATCTCATCCCTGCTGATACGG 2100
Qy 266 CysTyrSerLysAspProSerGlnArgProSerMetGluGluIleVal 281
Db 2101 GCGTGAATGATCATGCTCTGAGAGCAGCAGCATTCCTCTGAAGTCGTT 2148

RESULT 11
US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN1-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48) ..(2552)
US-09-947-199-1

Alignment Scores:
Pred. No.: 2,34e-38 length: 3025
Score: 410.00 Matches: 103
```

```
Percent Similarity: 57.89% Conservative: 51
Best Local Similarity: 38.72% Mismatches: 91
Query Match: 25.39% Indels: 22
DB: 10 Gaps: 9

US-09-830-144-2_COPY_1_303 (1-303) x US-09-947-199-1 (1-3025)

Qy 31 IleAspTyrLysGluIleGluValGluGluValAlaGlyArgGlyAlaPheGlyValVal 50
Db 1419 CTTACGCTCTCAGAAATTTGAGTTCCATGAGATTAATGGCTTCAGAGTTCTTTGGGAAGTA 1478
Qy 51 CysLysAlaLysTyrPAlaLysAspValAlaIleLysGlnIleGlu----- 66
Db 1479 TATTAAGACGATCAGAAATTAATTAATGTGGCTATTAACAGTTATGAGCCAAATACCTTAC 1538
Qy 67 ---SerGluSerGluArgLysAlaPheIleValGluLeuArgGlnLeuSerArgValAsn 85
Db 1539 TGCTCCAAAGTCAGATGAGATGATGTTTGGCCGAGAGGTGTCATTCCTGCGCAGCTCAAT 1598
Qy 86 HisProAsnIleValLysLeuTyrGlyAlaCysLeuAsn-----ProValCysLeu 102
Db 1599 CATCCCTGCGTAATTCAGTTGTGGGTGCTGCTGATGATCCAGCCAGTTTGGCATTT 1658
Qy 103 ValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHISGlyAlaGluProLeu 122
Db 1659 GTCACTCAATACATATGAGGAGGTTCCTCTCTCCTCTTATAGA-GCAGAGAGAT 1717
Qy 123 ProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeuGlnCysSerGln-GlyValAl 142
Db 1718 TCTTGATTT-----GCAGCTAAATTAATTAATTTGAGTGAATGTTGCCAAGGATGGA 1771
Qy 142 aTyrLeuHisSerMet--GlnProLysAlaLeuIleHisArgAspLeuLysProProAs 161
Db 1772 GTACCTTCACAACTGACACAGCCA-----ATTATACATCGTGAACCTTGMAACGTACAA 1825
Qy 161 nLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAs 181
Db 1826 TATCTCTCTTATGAGATGAGGATGCTGTG---GTGGCAGATTTTGGAAATCAAGATT 1882
Qy 181 PileGlnThr-----HisMetThrAsnAsnLysGlySerAlaAlaTyrMetAl 197
Db 1883 TCTACAGTCTGTGGATAAGAACATGACAAACAAACCTGGAACTCCGTTGGATGG 1942
Qy 197 aProGluValPhe--GluGlySerAsnTyrSerGluLysCysAspValPheSerTyrG 216
Db 1943 TCCGAGAGTGTTCACGACGAGCACTCGGTACACCATCAAGAGATGCTTCAGCTATGC 2002
Qy 216 YIleIleLeuTyrGluValIleThrArgLysProPheAspGluIleGlyProAl 236
Db 2003 TCTGTCTGTGGAAATTCCTCACTGGCGAAATTCATTCGCTCATCTCAAGCCAGCGG 2062
Qy 236 aPheArgIleMetTyrAlaHisAsnGlyThrArgProProLeuIleLysAsnLeuP 256
Db 2063 TGCGGACAGACATGGCTTACACACC---ATCAGACCTCCATGGGCTATTCATTC 2119
Qy 256 OlyAspProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 276
Db 2120 CAAGCCCATATCATCTCTGCTGATACGAGGTGAACGCATGTCTGAAGAGAACCCGA 2179
Qy 276 rMetGluGluIleVal 281
Db 2180 ATTTTCTGAAGTTGTC 2195

RESULT 12
US-09-947-199-7
; Sequence 7, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jayaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: MN1-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
```


QY 126 ThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHis 145
DB 2146 GATGAACACCGTCGAAATTAATATGCTTTGATGTGGCAAGGAAATGAATCACTACCCAC 2205
QY 146 SerMetGlnProLysAlaLeuLeuHisAspAspLeuLysProProAsnLeuLeuVal 165
DB 2206 AGACGATGATCT---CGAATGTTCATCGGATTTAAATTCACCGAATGTGTAGTTGAC 2262
QY 166 AlaGlyGlyThrValLeuLysIleCysAspPheGly---ThrAlaCysAspIleGlnThr 184
DB 2263 AAGACATATACAGTC---AAGGTTTGTGATTTGTCTCTCCCGTTTAAAGCAGCACA 2319
QY 185 HisMetThrAsnAsnLys-----GlySerAlaAlaTrpMetAlaProGluValPheGlu 202
DB 2320 TTTCTTTCATCCAAATCTGCAGCTGCAGCACTGAAATGATGACACAGAACTACAGCC 2379
QY 203 GlySerAsnTyrSerGlnLysCysAspValPheSerTrpGlyIleIleLeuTrpGluVal 222
DB 2380 GATGAACATCAATGAATAAGTCAGATGTTTACAGCTTGGATGATTTTGTGGAGTTG 2439
QY 223 IleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAla 242
DB 2440 GCAACTTTCGCAACGACGATGTGTAATCTA---AACCCAGCTCAGGTGTGCCAGCTGT 2496
QY 243 ValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeu 262
DB 2497 GGATTTAAGGCGAAAGGCTTGACATCCACGATGATTAATCCCAATGCTTCCTTA 2556
QY 263 MetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 282
DB 2557 ATAGTGCGCTTGCTGGCCGATGAGCATGAAACGTCCTTTTCCAGCATTTATGGA 2616
QY 283 IleMetThrHisLeuMetArgTyrPhePro 292
DB 2617 ACCTTGAAACCAATGACTAAACAGCGCA 2646

RESULT 14
US-09-938-842A-1577
Sequence 1577, Application US/0993842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SAME, AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 1577
LENGTH: 2211
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-1577

Alignment Scores:
Pred. No.: 2,22e-36 Length: 2211
Score: 391.50 Matches: 98
Percent Similarity: 53.36% Conservative: 53
Best Local Similarity: 34.63% Mismatches: 99
Query Match: 24.24% Indels: 33
DB: 9 Gaps: 13

US-09-830-144-2_COPY_1_303 (1-303) X US-09-938-842A-1577 (1-2211)

QY 30 GluIleAspTyrLysGluIleGluValGluGluValAlaGlyTrpGlyAlaPheGlyVal 49
DB 1384 GAGATACGATGGGAAAGATCTTACAACTGGGAGAGGTCCGAAGAGATTATTTGCTCG 1443
QY 50 ValCysLysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSer 69
DB 1444 GTTCATGCTGGAGTTTGGAAATGATCGGATGCTGTCTATTAAAG---GTTTACTTCGATGG 1500
QY 70 GluArgLysAlaPheIleVal-----GluLeuArgGlnLeuSerArgVal 84
DB 1501 GATTACATGAGATGATCTTACGAGAGTCGAAAGAGATCAACATTATGAAGAACTG 1560
QY 85 AsnHisProAsnIleValLysLeuTyrGlyValAcysLeuAsnProValCys-----101
DB 1561 AGACATCCGAATGTGCTACTATTATGGAGCA-----GTATGTACAGAGAA 1608
QY 102 -----LeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValIleHisGly 118
DB 1609 AAATCTGCCATTAATCAAGATATATATGCCAAGAGGAGTCTTCCAAATCTTCAATAT 1668
QY 119 Ala---GluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
DB 1669 ACGAATCAGCCATG-----GACAAAGAAAGCCGTTTAAATGAGCCCTTGATGT 1719
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuLeuHisArgAspLeu 157
DB 1720 GCTAGGGAATGAAATTACTTACACCGCAGAAATCC---CCAATTGTACTATAGACTTG 1776
QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
DB 1777 AAATCTTCATCTACTCTCGTGACAGCAAGAACTGCAATGTC---AAGTTGAGACTTTGG 1833
QY 178 -----ThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySer 192
DB 1834 TTATCAAGTGAAGAAAGCAACCTTCTTGATGACT-----AAATCCGGAAGAACT 1887
QY 193 AlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspVal 212
DB 1888 CCGCAGTGAATGCTCTGAGTTCTCAGAAATGAACCTTCGATGAAAGATGTGATG 1947
QY 213 PheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIle 232
DB 1948 TTCAGCTTGGAGTCATCTTATGAGAGCTATGACTACCTTATACCATGAGCCGTTG 2007
QY 233 GlyGlyProAlaPheArgIleMetTrpAlaVal---HisAsnGlyThrArgProProLeu 251
DB 2008 AAC-----TCTATTCAAGTTGTTGGAGTTGTTGATGATGACGATTAAGACTTA 2061
QY 252 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspPro 271
DB 2062 CTTGAAGATTAATCCCGGATCGCATCCATATACAGAGATTGTGGCAACTGATCCA 2121
QY 272 SerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPhe 291
DB 2122 GCAAAACGACGCTGCTCGAGGAATTAATATGACATATATGACGCTGTTCCGCAAA--- 2178
QY 292 ProGlyAla 294
DB 2179 CCAAGGTCA 2187

RESULT 15
US-09-938-842A-882
Sequence 882, Application US/0993842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SAME, AND METHODS OF USE
CURRENT APPLICATION NUMBER: US/09/938,842A

; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 882
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-882

Alignment Scores:
Pred. No.: 3,89e-36 Length: 1428
Score: 387.00 Matches: 96
Percent Similarity: 52.20% Conservative: 58
Best Local Similarity: 32.54% Mismatches: 109
Query Match: 23.96% Indels: 32
DB: 9 Gaps: 10

US-09-830-144-2_COPY_1_303 (1-303) x US-09-938-842A-882 (1-1428)

```
QY 10 SerSerSerSerSerAlaGlyGluMetIleGluAlaProSerGlnValLeuAsnPheGlu 29
DB 427 TCTAAGTAAGTTCTGCAGAGTGTCTTGAAGATGC-----462
QY 30 GluIleAspTyrLysGluIleGluValGluGluValValGlyArgGlyAlaPheGlyVal 49
DB 463 CTCATTGATGTCTTAAGTTGTCTTATGGGATAGTTTGTCTCACGGGAAATATAGCCAG 522
QY 50 ValCysLysAlaLysTrpArgAlaLysAspValAlaIleLysGlnIleGluSerGluSer 69
DB 523 ATTTATCATGGTGAATATGAAGCAAGCTGTGTCTCTGAAGATTATCACAGCGCTGAG 582
QY 70 Glu-----ArgLysAlaPheIleValGluLeuArg 79
DB 583 GATAGTGACGACATATCTTGGGAGCTCGTTTAGAAAAAGAGTTTATCGTGGAGCCACT 642
QY 80 GlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCysLeuAsnPro 99
DB 643 CTTTATCTCGACTAAGCCATCCAAATGCTGTTAAGTTTGTGGAGTGAATACTCGAAAC 702
QY 100 ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnValLeuHisGlyAla 119
DB 703 ---TGTCATCATCAGAGTATGTACCTCGAGGTCTTTAAGATCATATCTGCACAAAGCTC 759
QY 120 Glu-----ProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeuGlnCys 137
DB 760 GAGCAGAAATCCCTTCCT-----TTGGAACAGCTAATCGAATTTGGTCTGGATATT 810
QY 138 SerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeu 157
DB 811 GCTAAAGGAATGGAATATATTCACCTCA-----AGAGAGATAGTTTCATCAGGATCTG 861
QY 158 LysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGly 177
DB 862 AAGCCAGAAAC---GTGTTGATCGACAATGACTTTTCATTTGAAGATTCTGCACITTTGC 918
QY 178 ThrAlaCysAspIleGln-----ThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 195
DB 919 ATAGCGTGCAGGAGGAGTACTGTGATGTTTGGGGGATAACATAGAACTTATAGGTGG 978
QY 196 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 215
DB 979 ATGGCACCCTGAAGTTTAAACCGGATACCACATCGAGCGGAAGTGCATGTTTATAGTTT 1038
QY 216 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 235
DB 1039 GGACTCTTTTATGGGAAATGGGTAGCTGGAGCACTTCCATATGAGGAGATGAATTT--- 1095
QY 236 AlaPheArgIleMetTrpAlaVal---HisAsnGlyThrArgProLeuIleLysAsn 254
```

```
DB 1096 GCTGAACAAATTGCCTACGCAGTTATATACAAGAAAATTAGCCAGTTATACCGACGGAT 1155
QY 255 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 274
DB 1156 TGTCCAGCGGCCCATGAAGAGCTGATCGAGCGATGTTGGTCATCGCAACACACAAGAGA 1215
QY 275 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArg 289
DB 1216 CCGGAATTCTGGCAGATTGTCAAAGTGTGGAACATTTCAAGAAG 1260
```

Search completed: December 10, 2002, 03:46:27
Job time : 114 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:48:55 ; Search time 37 Seconds
(without alignments)
821.112 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252
Sequence: 1 VELRQLSRVNHPIVKLYCA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*

- 1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT.*
- 2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT.*
- 3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT.*
- 4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT.*
- 5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT.*
- 6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT.*
- 7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT.*
- 8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT.*
- 9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT.*
- 10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT.*
- 11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT.*
- 12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT.*
- 13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT.*
- 14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT.*
- 15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT.*
- 16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT.*
- 17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT.*
- 18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT.*
- 19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT.*
- 20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT.*
- 21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT.*
- 22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT.*
- 23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	567	20 AAY28998	Human TGF-beta act
2	1252	100.0	579	18 AAW27092	Mouse transforming
3	1252	100.0	579	18 AAW27093	Human transforming
4	1252	100.0	579	20 AAY28996	Human TGF-beta act
5	1252	100.0	579	20 AAY09542	Human TAK1 protein
6	1252	100.0	579	21 AAY91000	Human TAK-1 protei
7	1252	100.0	579	23 ABB85033	Pain regulated pro
8	1252	100.0	590	20 AAY09547	Human TAK1-6xHis p
9	1252	100.0	606	20 AAY28997	Human TGF-beta act
10	700	55.9	678	22 ABB58061	Drosophila melanog

11	414	33.1	252	22	ABB60985	Drosophila melanog
12	374	29.9	367	21	AAG32053	Arabidopsis thalia
13	374	29.9	369	21	AAG32052	Arabidopsis thalia
14	374	29.9	407	21	AAG32051	Arabidopsis thalia
15	373	29.8	369	21	AAG22172	Arabidopsis thalia
16	373	29.8	374	21	AAG22171	Arabidopsis thalia
17	373	29.8	412	21	AAG22170	Arabidopsis thalia
18	371	29.6	349	22	AAG75571	Human colon cancer
19	371	29.6	455	21	AA18657	A human regulator
20	371	29.6	455	21	AY833278	Human survival reg
21	371	29.6	455	21	AY843321	A human cardiovasc
22	371	29.6	473	22	AAW25322	Human protein sequ
23	371	29.6	800	22	AAW71957	Human TGF-beta rec
24	371	29.6	800	22	AAB65673	Novel protein kina
25	363.5	29.0	1020	22	ABBS8999	Drosophila melanog
26	355.5	28.4	1021	23	ABP61000	Novel human protei
27	355	28.4	719	22	AA885513	Human protein kina
28	355	28.4	1036	23	ABB80923	Novel human protei
29	353.5	28.2	859	16	AAW82886	Human leucine zipp
30	353.5	28.2	859	18	AAW31127	Human leucine-zipp
31	352.5	28.2	888	23	ABBS7049	Mouse ischaemic co
32	351.5	28.1	977	22	ABW71694	Drosophila melanog
33	350	28.0	589	21	AA453984	Arabidopsis thalia
34	350	28.0	732	21	AA453983	Arabidopsis thalia
35	350	28.0	760	21	AA453982	Arabidopsis thalia
36	349.5	27.9	1046	22	AAE11775	Human kinase (PKIN
37	349.5	27.9	1097	23	AAE21717	Human PKIN-12 prot
38	348	27.8	319	21	AAQ25601	Arabidopsis thalia
39	348	27.8	341	21	AAQ25600	Arabidopsis thalia
40	348	27.8	391	21	AAQ25599	Arabidopsis thalia
41	347	27.7	338	21	AAQ28422	Arabidopsis thalia
42	347	27.7	338	21	AAQ50301	Arabidopsis thalia
43	347	27.7	346	21	AAQ28421	Arabidopsis thalia
44	347	27.7	346	21	AAQ50300	Arabidopsis thalia
45	345.5	27.6	847	23	AAE22763	Human mitogen acti

ALIGNMENTS

RESULT 1

AAAY28998

ID AAY28998 standard; Protein; 567 AA.

XX

AC AAY28998;

XX

DT 29-OCT-1999 (first entry)

XX

Human TGF-beta activated kinase (TAK) 1c amino acid sequence.

XX

Nuclear factor kappa B; NF-kB, inhibitor; TGF-beta activated kinase 1;
TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
intracrabile disease; atrophic dermatitis; psoriasis; viral infection;
endotoxin shock; septicemia; human; hTAK1c.

XX

OS Homo sapiens.

XX

PN WO9940202-A1.

XX

PD 12-AUG-1999.

XX

PF 02-FEB-1999; 99WO-JP00422.

XX

PR 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

XX

PA (TANA) TANABE SEIVAKU CO.

XX

PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX WPI; 1999-494298/41.

DR N-PSDB; AAX99698.

XX

PT Nuclear factor kappa B activation inhibitors, useful as preventives
PT For, e.g. autoimmune diseases
PS Examples: Page 43-46; 49pp; Japanese.
XX
CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1c
CC (hTAK1c) protein.
CC
SQ Sequence 567 AA;
Query Match 100.0%; Score 1252; DB 20; Length 567;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELRQLSRVNHPNIVKLYGACLNPCVLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPNIVKLYGACLNPCVLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLKRPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 136 QCSQGVAYLHSMQPKALIHRLKRPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSMGIIWMEVITRRKPFDEIGGPAFRIIMAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSGNSYSEKCDVFSMGIIWMEVITRRKPFDEIGGPAFRIIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFGADEPLQYPCQ 303
RESULT 2
AAW27092
ID AAW27092 standard; Protein; 579 AA.
XX
AC AAW27092;
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1.
XX
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Mus musculus.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
XX
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
XX
PA (UENO/) UENO N.
XX
DR WPI; 1997-380171/35.
XX
DR N-PSDB; AAT85094.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
XX
PT - useful for studying the TGF-beta signal transduction system
XX
PS Claim 14; Page 10-12; 20pp; Japanese.

XX
CC The present sequence represents mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
CC
SQ Sequence 579 AA;
Query Match 100.0%; Score 1252; DB 18; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELRQLSRVNHPNIVKLYGACLNPCVLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHPNIVKLYGACLNPCVLWMEYAGSLYVNLHGAEPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIHRLKRPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
DB 136 QCSQGVAYLHSMQPKALIHRLKRPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSMGIIWMEVITRRKPFDEIGGPAFRIIMAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSGNSYSEKCDVFSMGIIWMEVITRRKPFDEIGGPAFRIIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPSMEIIVKIMTHLMRYFGADEPLQYPCQ 303
RESULT 3
AAW27093
ID AAW27093 standard; Protein; 579 AA.
XX
AC AAW27093;
XX
DT 19-NOV-1997 (first entry)
XX
DE Human transforming growth factor-beta activated kinase TAK-1.
XX
KW TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase.
XX
OS Homo sapiens.
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0685625.
XX
PR 29-SEP-1995; 95JP-0253549.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
XX
PA (UENO/) UENO N.
XX
DR WPI; 1997-380171/35.
XX
DR N-PSDB; AAT85095.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
XX
PT - useful for studying the TGF-beta signal transduction system
XX
PS Claim 15; Page 13-15; 20pp; Japanese.
XX
CC The present sequence represents human transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone

CC morphogenetic protein (BMP) and activates MAPK kinase by
 XX phosphorylation.

SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 18; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60
 DB 76 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135
 QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
 DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
 QY 121 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
 DB 196 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
 QY 181 PKPIESLMTRCWSKDPQSRPSMEIEIVKIMTHLMRYFPGADEPLOYPCQ 228
 DB 256 PKPIESLMTRCWSKDPQSRPSMEIEIVKIMTHLMRYFPGADEPLOYPCQ 303

RESULT 4

AAAY28996
 ID AAY28996 standard; Protein; 579 AA.

AC AAY28996;

DT 29-OCT-1999 (first entry)

DE Human TGF-beta activated kinase (TAK) 1a amino acid sequence.

KW Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
 KW TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
 KW intractable disease; atrophic dermatitis; psoriasis; viral infection;
 KW endotoxin shock; septicemia; human; hTAK1a.

OS Homo sapiens.

PN WO9940202-A1.

PD 12-AUG-1999.

PF 02-FEB-1999; 99WO-JP00422.

PR 30-OCT-1998; 98JP-0309316.

PR 06-FEB-1998; 98JP-0026003.

XX (TANA) TANABE SEIYAKU CO.

XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;

XX WPI; 1999-494298/41.

DR N-PSDB; AAX99696.

PT Nuclear factor kappa B activation inhibitors, useful as preventives
 PT for, e.g. autoimmune diseases

PS Examples; Page 35-39; 49pp; Japanese.

XX The invention provides a method for identifying or screening a nuclear
 CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
 CC test substance on modulating the function(s) of TGF-beta activated kinase
 CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
 CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
 CC intractable diseases with inflammation (such as atrophic dermatitis and
 CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
 CC present sequence represents the amino acid sequence of human TAK1a
 CC (hTAK1a) protein.

XX SQ Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;
 Best Local Similarity 100.0%; Pred. No. 2.9e-127;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 60
 DB 76 VELRQLSRVNHNPNIKLYGACLNPCVCLVMEYAEAGGSLYNVLHGAELPPYYTAHAHMSWCL 135
 QY 61 QCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
 DB 136 QCSQGVAYLHSMOPKALIHRLDKPPNLLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
 QY 121 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
 DB 196 MAPEVFEFGSNYSEKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
 QY 181 PKPIESLMTRCWSKDPQSRPSMEIEIVKIMTHLMRYFPGADEPLOYPCQ 228
 DB 256 PKPIESLMTRCWSKDPQSRPSMEIEIVKIMTHLMRYFPGADEPLOYPCQ 303

RESULT 5

AAAY09542

ID AAY09542 standard; Protein; 579 AA.

AC AAY09542;

DT 21-JUL-1999 (first entry)

DE Human TAK1 protein.

KW Human; TAK1; TAK1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 183..1922

FT /*tag= a

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

XX N-PSDB; AAX56279.

PT Screening for TGF- beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder

PS Claim 4; Page 155-157; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAK1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAK1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors or activators, or cell proliferation prevention inhibitors or
 CC activators, or monocyte migration inhibitors or activators, or

CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents human TAK1.

XX Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 20; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWCL 60
DB 76 VELRLSRVNHPIVLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIOHTMTNKGSAAM 120
DB 136 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIOHTMTNKGSAAM 195
QY 121 MAPEVFECSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 180
DB 196 MAPEVFECSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 6
AA91000
ID AA91000 standard; Protein; 579 AA.

XX AA91000;

DT 04-SEP-2000 (first entry)

DE Human TAK-1 protein sequence SEQ ID NO:2.

XX Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
XX screening; signal transduction; inhibition; inflammatory cytokine;
XX IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
XX antiinflammatory; suppression.

OS Homo sapiens.

PN WO200023610-A1.

PD 27-APR-2000.

PF 21-OCT-1999; 99WO-JP05817.

PR 21-OCT-1998; 98JP-0299962.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;

DR WPI; 2000-339707/29.

DR N-PSDB; AAA39105.

XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents -

PS Example 1; Page 80-84; 100pp; Japanese.

XX The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAb1 and selecting for inhibition
CC of TAK1/TAb1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug

CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence represents human TAK-1, which is used in the exemplification of
CC the present invention.

XX Sequence 579 AA;

Query Match 100.0%; Score 1252; DB 21; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWCL 60
DB 76 VELRLSRVNHPIVLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAHAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIOHTMTNKGSAAM 120
DB 136 QCSQGVAYLHSMQPKALLHRDLKPPNLLVAGTVLKICDFGTACDIOHTMTNKGSAAM 195
QY 121 MAPEVFECSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 180
DB 196 MAPEVFECSNYSKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPQSPMEIYKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 7

ABB85033
ID ABB85033 standard; Protein; 579 AA.

XX ABB85033;

DT 16-MAY-2002 (first entry)

DE Pain regulated protein sequence 28.

XX Pain; analgesic; gene therapy; neurological disorder;
XX neurodegenerative disease.

OS Homo sapiens.

PN WO200212338-A2.

PD 14-FEB-2002.

PF 03-AUG-2001; 2001WO-EP09011.

PR 03-AUG-2000; 2000DE-1037759.

PA (CHEF) GRUENENTHAL GMBH.

PI Gillen C, Wetzel S, Wendt S, weine E, Schaefer MK;

DR WPI; 2002-257469/30.

DR N-PSDB; ABL88437.

XX Identifying pain-regulating compounds, useful for treating chronic pain
PT and for diagnosis, by measuring binding of compounds to specific
PT peptides and proteins -

PS Claim 1; Fig 44; 213pp; German.

XX The invention relates to identifying pain-regulating substances (A)
CC comprising (i) incubating a test substance with a cell (or preparation
CC from it) that has synthesised a peptide or protein (B) and (ii) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying

pain-regulating substances (A) with analgesic activity, (A) along with nucleic acid (AB188441-2BL88441) that encode proteins (B) AB85006-AB85037) that interact with (A); (B): vectors containing the nucleic acid; antibodies against (B); cells that express (B) and agents that bind to (B), are all useful for treating pain, particularly chronic pain, including use in gene therapy. The same materials can also be used for diagnosis, e.g. of neurological and neurodegenerative diseases. The present sequence is that of a polypeptide of the invention.

```
Query Match      100.0%; Score 1252; DB 23; Length 579;
Best Local Similarity 100.0%; Pred. NO. 2.9e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1	VEURQLSRVNHPIV	KLYGACLN	PVCLVMEY	ABGSGSYNV	LHGA	BELPY	YTTAAH	AMSWCL	60
Db	76	VEURQLSRVNHPIV	KLYGACLN	PVCLVMEY	ABGSGSYNV	LHGA	BELPY	YTTAAH	AMSWCL	135
Qy	61	QCSOGVAYLHSMOP	KALIH	RDLP	PNLLVAGGT	VLKICD	FGTACDI	QTHMTN	KGSAAW	120
Db	136	QCSOGVAYLHSMOP	KALIH	RDLP	PNLLVAGGT	VLKICD	FGTACDI	QTHMTN	KGSAAW	195
Qy	121	MAPEVFGSGSYN	SEKCDV	FSWGII	LWEVIT	TRRKPDE	IGGPAFRIM	WAVHNG	TRPPLIKNL	180
Db	196	MAPEVFGSGSYN	SEKCDV	FSWGII	LWEVIT	TRRKPDE	IGGPAFRIM	WAVHNG	TRPPLIKNL	255
Qy	181	PKPTESLMTWC	SKDPSQ	PSMSEIV	KIMTHLM	RYPPGA	DEPLOY	PCQ	228	
Db	256	PKPTESLMTWC	SKDPSQ	PSMSEIV	KIMTHLM	RYPPGA	DEPLOY	PCQ	303	

RESULT 8
AAY09547
ID AAY09547 standard: Protein: 590 AA:

XX	AA09547;
AC	21-JUL-1999 (first entry)
XX	Human TAK1-6xHis protein.
DT	Human; TAB1; TAK1; screening; inhibition; TGF-beta;
DT	transforming growth factor beta.
XX	
XX	
DE	
XX	
KW	
KW	

XX	OS	Homo sapiens.
XX	OS	Synthetic.
XX	PN	WO9921010-A1.
XX	XX	
XX	PD	29-APR-1999.
XX	XX	
XX	PF	22-OCT-1998; 98WO-JF04796.
XX	XX	
XX	PR	22-OCT-1997; 97JP-0290188.
XX	XX	
XX	PA	(CHUS) CHUGAI SEIYAKU KK.
XX	XX	
XX	FI	Ohtomo T, Ono K, Tsuchiya M;
XX	XX	WPI; 1999-312645/26.
XX	DR	N-PSDB: AAX56285.
XX	DR	

XX Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder
XX
PS Example 1; Page 171-174; 195pp; Japanese.

AA A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC

CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents TAK1-6xHis from an example of
CC the present invention.

Query Match 100.0%; Score 1252; DB 20; Length 590;
Best Local Similarity 100.0%; Pred. No. 3e-127;

Qy	1	VELRQLSRVNHPIVKLYGACLNPCVLWMEYABEGGSLYNVLHGAEPLPYTTAAHANSWCL	60
Db	76	VELRQLSRVNHPIVKLYGACLNPCVLWMEYABEGGSLYNVLHGAEPLPYTTAAHANSWCL	135
Qy	61	QCSGGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGFACDIQTHMTNNGKSAAM	120
Db	136	QCSGGVAYLHSMOPKALIHRLDKPPNLLVAGGTVLKICDGFACDIQTHMTNNGKSAAM	195
Qy	121	MAPEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMVAVHNGTRPPLIKNL	180
Db	196	MAPEVFEGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMVAVHNGTRPPLIKNL	255
Qy	181	PKPIESLMTKCSKDPQRSMSMEIVKIMTHLMRYPPGADEPLQYPCQ	228
Db	256	PKPIESLMTKCSKDPQRSMSMEIVKIMTHLMRYPPGADEPLQYPCQ	303

RESULT 9
AAY28997
ID AAY28997 standard: Protein: 606 AA.

XX	AA28997;
AC	
XX	
XX	29-OCT-1999 (first entry)
DT	
XX	
XX	Human TGF-beta activated kinase (TAK) 1b amino acid sequence.
DE	
XX	
KW	Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated
KW	TAK1; autoimmune disease; chronic rheumatoid arthritis; infla-
KW	tractable disease; atrophic dermatitis; psoriasis; viral in-
KW	endotoxin shock; septicemia; human; IRAK1b.

XX	Homo sapiens.	
OS		
XX		
XX	W09940202-A1.	
FN		
XX		
XX	12-AUG-1999.	
PD		
XX		
XX	02-FEB-1999;	99WO-JP00422.
PF		
XX		
PR	30-OCT-1998;	98JP-0309316.
PR	06-FEB-1998;	98JP-0026003.
PR		
XX	(TANA) TANABE SEIYAKU CO.	
PA		

PI Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX
XX WPI; 1999-494298/41.
DR N-PSDB; AAX99637.
XX
XX
PT Nuclear factor kappa B activation inhibitors, useful as preventives
PT for, e.g. autoimmune diseases
XX
PS Examples; Page 39-43; 49pp; Japanese.

CC The invention provides a method for identifying or screening a nuclear
CC factor kappa B (NF-kB) activation inhibitor by examining the effect of a
CC test substance on modulating the function(s) of TGF-beta activated kinase
CC 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
CC treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
CC intractable diseases with inflammation (such as atrophic dermatitis and
CC psoriasis), viral infection, endotoxin shock, septicemia and others. The
CC present sequence represents the amino acid sequence of human TAK1b
CC (hTAK1b) protein.
CC
SQ Sequence 606 AA;
Query Match 100.0%; Score 1252; DB 20; Length 606;
Best Local Similarity 100.0%; Pred. No. 3,1e-127;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELRLQSRVNHNPVIVKLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAAHMSMCL 60
DB 76 VELRLQSRVNHNPVIVKLYGACINPVCLVMEYAGGSLYVNLHGAEPLPYTAAHMSMCL 135
QY 61 QCSQGVAYIHSNQPKALIRDLKRPVLLVAGTVLKITCDFTACDICTHMTNNGSAAW 120
DB 136 QCSQGVAYIHSNQPKALIRDLKRPVLLVAGTVLKITCDFTACDICTHMTNNGSAAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGSPAFRIMWAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGSPAFRIMWAVHNGTRPPLIKNL 255
QY 181 KPRIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFFGADDEPLQYPCQ 228
DB 256 KPRIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFFGADDEPLQYPCQ 303
RESULT 10
ABB58061
ID ABB58061 standard; Protein; 678 AA.
XX
AC ABB58061;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 975.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI; 2001-656860/75.
XX
DR N-PSDB; ABL02164.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 975; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ffp.wipo.int/pub/published_pct_sequences.
CC
SQ Sequence 678 AA;
Query Match 55.9%; Score 700; DB 22; Length 678;
Best Local Similarity 57.5%; Pred. No. 4,9e-67;
Matches 130; Conservative 35; Mismatches 57; Indels 4; Gaps 3;
QY 2 ELRLQSRVNHNPVIVKLYG--ACINPVCLVMEYAGGSLYVNLHGAEPLPYTAAHMSMC 59
DB 60 EVKQLSRVKNPIIALHGISYQOATYLIWFPAEGSLNHFHNG-KVKPAVSLAHMSMA 118
QY 60 LQCSQGVAYIHSNQPKALIRDLKRPVLLVAGTVLKITCDFTACDICTHMTNNGSAA 119
DB 119 RQCEGLAYIHAHTPKPLIRDVKPLNLLITNKGRLKICDFETVADKSTMTNNGSAA 178
QY 120 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGSPAFRIMWAVHNGTRPPLIKN 179
DB 179 MAPEVEGSGNSYSEKCDVFSMGIIIMEVITRRKPFDEIGSPAFRIMWAVHNGTRPPLIKN 237
QY 180 KPRIESLMTRCWSKDPSPQSPSMEIEIVKIMTHLMRYFFGADDEPLQY 225
DB 238 CPKRIEDLMTACWKVYEDPSPQSPSMEIEIVKIMTHLMRYFFGADDEPLQY 283
RESULT 11
ABB60985
ID ABB60985 standard; Protein; 252 AA.
XX
AC ABB60985;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster polypeptide SEQ ID NO 974.
XX
KM Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI; 2001-656860/75.
XX
DR N-PSDB; ABL05088.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Disclosure; SEQ ID NO 9747; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of

PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136031.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157753.
PR 07-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.

```
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.
```

Query Match 29.9%; Score 374; DB 21; Length 369;

Best Local Similarity 39.4%; Pred. No. 6, 5e-32; Mismatches 70; Indels 22; Gaps 9;

Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

```
QY 2 ELRQLSRVNHPIVVKLYGACLNLPV--CLVMEYAEGSGLYNVL----HGAEPLPYTAHA 55
DB 137 EYVSMNLAKHPINIVRFICGCRKPMWCVITEYAKGGSVHQFLTRQRNRAVPLKL-----A 191
QY 56 MSNCLQSCGQVATVHSMQKALIHRLKRPNIILVAGTVLKCIPGRTA-CDIQTH-MTN 113
DB 192 VKQALDVAGMAVHG--RNFLHRLKSDNLLISADKSI-KIADGVARIIVQTEGMRP 247
QY 114 NKGSAAWMAPEVPEGSNYSKCDVFSWGIILWEVITRRKPFDEIGS--PAFRIMAVHNG 171
DB 248 ECTTYMMAPEMIQHRAYNQKDVYSFGIVLWELITGLLPQNMRAVQAFAFV--VVRG 304
QY 172 TRPPLIKNLPKPIESLMTKWSKDSQSPSMEEIVKIM 209
DB 305 VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVEVVKLL 342
```

RESULT 14

AAG32051

ID AAG32051 standard; Protein; 407 AA.

XX AAG32051;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 38595.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

XX Arabidopsis thaliana.

OS EPI033405-A2.

PN 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

```
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 11-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139899.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140981.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
```

PR 19-JUL-1999;	99US-0144335.	PR 14-OCT-1999;	99US-0159330.
PR 20-JUL-1999;	99US-0144352.	PR 14-OCT-1999;	99US-0159331.
PR 20-JUL-1999;	99US-0144632.	PR 14-OCT-1999;	99US-0159637.
PR 20-JUL-1999;	99US-0144884.	PR 14-OCT-1999;	99US-0159638.
PR 21-JUL-1999;	99US-0144814.	PR 18-OCT-1999;	99US-0159584.
PR 21-JUL-1999;	99US-0145086.	PR 21-OCT-1999;	99US-0160741.
PR 21-JUL-1999;	99US-0145088.	PR 21-OCT-1999;	99US-0160767.
PR 22-JUL-1999;	99US-0145085.	PR 21-OCT-1999;	99US-0160768.
PR 22-JUL-1999;	99US-0145087.	PR 21-OCT-1999;	99US-0160770.
PR 22-JUL-1999;	99US-0145089.	PR 21-OCT-1999;	99US-0160814.
PR 22-JUL-1999;	99US-0145192.	PR 21-OCT-1999;	99US-0160815.
PR 23-JUL-1999;	99US-0145118.	PR 22-OCT-1999;	99US-0160980.
PR 23-JUL-1999;	99US-0145214.	PR 22-OCT-1999;	99US-0160981.
PR 23-JUL-1999;	99US-0145224.	PR 22-OCT-1999;	99US-0160989.
PR 26-JUL-1999;	99US-0145276.	PR 25-OCT-1999;	99US-0161404.
PR 27-JUL-1999;	99US-0145913.	PR 25-OCT-1999;	99US-0161405.
PR 27-JUL-1999;	99US-0145918.	PR 25-OCT-1999;	99US-0161406.
PR 27-JUL-1999;	99US-0145919.	PR 26-OCT-1999;	99US-0161359.
PR 28-JUL-1999;	99US-0145951.	PR 26-OCT-1999;	99US-0161360.
PR 02-AUG-1999;	99US-0146386.	PR 26-OCT-1999;	99US-0161361.
PR 02-AUG-1999;	99US-0146388.	PR 28-OCT-1999;	99US-0161920.
PR 02-AUG-1999;	99US-0146389.	PR 28-OCT-1999;	99US-0161992.
PR 03-AUG-1999;	99US-0147038.	PR 28-OCT-1999;	99US-0161993.
PR 04-AUG-1999;	99US-0147204.	PR 29-OCT-1999;	99US-0162142.
PR 05-AUG-1999;	99US-0147302.		
PR 05-AUG-1999;	99US-0147192.		
PR 05-AUG-1999;	99US-0147260.		
PR 06-AUG-1999;	99US-0147303.		
PR 06-AUG-1999;	99US-0147416.		
PR 09-AUG-1999;	99US-0147493.		
PR 09-AUG-1999;	99US-0147935.		
PR 10-AUG-1999;	99US-0148171.		
PR 11-AUG-1999;	99US-0148319.		
PR 12-AUG-1999;	99US-0148341.		
PR 13-AUG-1999;	99US-0148565.		
PR 13-AUG-1999;	99US-0148684.		
PR 16-AUG-1999;	99US-0149368.		
PR 17-AUG-1999;	99US-0149175.		
PR 18-AUG-1999;	99US-0149426.		
PR 20-AUG-1999;	99US-0149722.		
PR 20-AUG-1999;	99US-0149723.		
PR 20-AUG-1999;	99US-0149929.		
PR 23-AUG-1999;	99US-0149902.		
PR 23-AUG-1999;	99US-0149930.		
PR 25-AUG-1999;	99US-0150566.		
PR 26-AUG-1999;	99US-0150884.		
PR 27-AUG-1999;	99US-0151065.		
PR 27-AUG-1999;	99US-0151066.		
PR 27-AUG-1999;	99US-0151080.		
PR 30-AUG-1999;	99US-0151303.		
PR 31-AUG-1999;	99US-0151438.		
PR 01-SEP-1999;	99US-0151930.		
PR 07-SEP-1999;	99US-0152363.		
PR 10-SEP-1999;	99US-0153070.		
PR 13-SEP-1999;	99US-0153758.		
PR 15-SEP-1999;	99US-0154018.		
PR 16-SEP-1999;	99US-0154039.		
PR 20-SEP-1999;	99US-0154779.		
PR 22-SEP-1999;	99US-0155139.		
PR 23-SEP-1999;	99US-0155486.		
PR 24-SEP-1999;	99US-0155659.		
PR 28-SEP-1999;	99US-0156458.		
PR 29-SEP-1999;	99US-0156596.		
PR 04-OCT-1999;	99US-0157117.		
PR 05-OCT-1999;	99US-0157753.		
PR 06-OCT-1999;	99US-0157865.		
PR 07-OCT-1999;	99US-0158029.		
PR 08-OCT-1999;	99US-0158232.		
PR 12-OCT-1999;	99US-0158369.		
PR 13-OCT-1999;	99US-0159293.		
PR 13-OCT-1999;	99US-0159294.		
PR 13-OCT-1999;	99US-0159295.		
PR 14-OCT-1999;	99US-0159329.		
<hr/>			
PR 14-OCT-1999;	99US-0159330.	PR 14-OCT-1999;	99US-0159330.
PR 14-OCT-1999;	99US-0159331.	PR 14-OCT-1999;	99US-0159331.
PR 14-OCT-1999;	99US-0159637.	PR 14-OCT-1999;	99US-0159637.
PR 14-OCT-1999;	99US-0159638.	PR 14-OCT-1999;	99US-0159638.
PR 18-OCT-1999;	99US-0159584.	PR 18-OCT-1999;	99US-0159584.
PR 21-OCT-1999;	99US-0160741.	PR 21-OCT-1999;	99US-0160741.
PR 21-OCT-1999;	99US-0160767.	PR 21-OCT-1999;	99US-0160767.
PR 21-OCT-1999;	99US-0160768.	PR 21-OCT-1999;	99US-0160768.
PR 21-OCT-1999;	99US-0160770.	PR 21-OCT-1999;	99US-0160770.
PR 21-OCT-1999;	99US-0160814.	PR 21-OCT-1999;	99US-0160814.
PR 21-OCT-1999;	99US-0160815.	PR 21-OCT-1999;	99US-0160815.
PR 22-OCT-1999;	99US-0160980.	PR 22-OCT-1999;	99US-0160980.
PR 22-OCT-1999;	99US-0160981.	PR 22-OCT-1999;	99US-0160981.
PR 22-OCT-1999;	99US-0160989.	PR 22-OCT-1999;	99US-0160989.
PR 25-OCT-1999;	99US-0161404.	PR 25-OCT-1999;	99US-0161404.
PR 25-OCT-1999;	99US-0161405.	PR 25-OCT-1999;	99US-0161405.
PR 25-OCT-1999;	99US-0161406.	PR 25-OCT-1999;	99US-0161406.
PR 26-OCT-1999;	99US-0161359.	PR 26-OCT-1999;	99US-0161359.
PR 26-OCT-1999;	99US-0161360.	PR 26-OCT-1999;	99US-0161360.
PR 26-OCT-1999;	99US-0161361.	PR 26-OCT-1999;	99US-0161361.
PR 28-OCT-1999;	99US-0161920.	PR 28-OCT-1999;	99US-0161920.
PR 28-OCT-1999;	99US-0161992.	PR 28-OCT-1999;	99US-0161992.
PR 28-OCT-1999;	99US-0161993.	PR 28-OCT-1999;	99US-0161993.
PR 29-OCT-1999;	99US-0162142.	PR 29-OCT-1999;	99US-0162142.
<hr/>			
Query Match 29.9%; Score 374; DB 21; Length 407;			
Best Local Similarity 39.4%; Pred. No. 7.4e-32;			
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;			
Qy 2 ELRLSRVHPNIVKLYGACLNIV--CLWMEYAEGLSLNVL-----HGAEPLEPYTAAHA 55			
Db 175 EVSMLANKHENIVRFITGACRKPVMWCIVTEYAKGGSVRQFLTRQNRVAPLKL-----A 229			
Qy 56 MSWCLQCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTA-CDIQTH-MTN 113			
Db 230 VKQALDVARGMAVHG---RNFHRLKSNLLISADKSI-KIADFGVARIEVQTEGMP 285			
Qy 114 NKGSAAMWAPFEGSNYSKCDVFSWGIILWEVITRRKPFDEIGG--PAPRIMWAVHNG 171			
Db 286 ETGYRWMAPEMIQHRAYNQKVDVSPGIVLWELITGLLPQNMTAVQAAFAV---VNRG 342			
Qy 172 TRPLIKNLKPFTESLMTRCWSKDPSPORPSMEIIVKIM 209			
Db 343 VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVGVKLL 380			
<hr/>			
RESULT 15			
ID AAG22172	AAG22172 standard; Protein; 369 AA.		
XX AC	AAG22172;		
XX DT	17-OCT-2000 (first entry)		
XX DE	Arabidopsis thaliana protein fragment SEQ ID NO: 24999.		
XX KW	Protein identification; signal transduction pathway; metabolic pathway;		
XX KW	hybridisation assay; genetic mapping; gene expression control; promoter;		
XX OS	Arabidopsis thaliana.		
XX PN	EP1033405-A2.		
XX PD	06-SEP-2000.		
XX PF	25-FEB-2000; 2000EP-0301439.		
XX PR	25-FEB-1999; 99US-0121825.		
PR 05-MAR-1999;	99US-0123180.		
PR 09-MAR-1999;	99US-0123548.		
PR 23-MAR-1999;	99US-0125788.		
PR 25-MAR-1999;	99US-0126264.		

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135253.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136382.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 26-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146386.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

```
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161922.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 29.8%; Score 373; DB 21; Length 369;
Best Local Similarity 39.9%; Pred. No. 8.3e-32;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHNPVIVKLYGACLNVP--CLVMEYAEAGSGLYNVL----HGAELPYTTAHA 55
Db 137 EVSMLAFLKHPNIVRFIGACIKPMVWCIVTEYAKGGSVRQFLTKRQRAVELKL-----A 191

QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTA-CDIOTH-MTN 113
Db 192 VMQALDVARGMAYVHE---RNFHRLKSDNLLISADRSI-KIADFGVARIETEGMTP 247

QY 114 NKGSAAMAPVEFGSNSEKCDVFSWGIILWEVITRRKPFDEIGG--PAPRIMAVHNG 171
Db 248 ETGTYRWMAPEMIQHRPPTYQKVDVYSFGIVLWELITGLLPFQNMATAVQAFAV---VNRG 304

QY 172 TRPPLIKNLPKPIESLMTFCWSKDPQRSMBEIVKIM 209
Db 305 VRPTVPADCLPVLGEIMTRCWDADPEVRPCFAEIVNLL 342
```

Search completed: December 10, 2002, 03:57:13
Job time : 38 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:52:45 ; Search time 20 Seconds
(without alignments)
1095.932 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELRQLSRVNHNPVVKLYGA.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 73: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1252	100.0	567	JC5957	transforming growt
2	1252	100.0	579	JC5955	transforming growt
3	1252	100.0	606	JC5956	transforming growt
4	392.5	31.3	954	S68178	mixed-lineage prot
5	374	29.9	407	G84635	probable protein k
6	373	29.8	412	T10671	protein kinase hom
7	363.5	29.0	394	JU0229	mixed-lineage prot
8	362	28.9	1338	T18287	protein-tyrosine k
9	356	28.4	462	S29851	protein kinase 6 (
10	355.5	28.4	370	T46150	protein kinase ATN
11	353.5	28.2	668	JC2363	protein kinase (BC
12	353.5	28.2	888	A53118	serine/threonine p
13	351.5	28.1	888	JC5399	dual leucine zippe
14	350	28.0	546	D84555	probable protein k
15	348	27.8	391	T48115	protein kinase ATN
16	347	27.7	475	T12955	probable protein k
17	346	27.6	328	T16747	hypothetical prote
18	346	27.6	1030	F96763	hypothetical prote
19	345.5	27.6	553	T04683	mixed-lineage prot
20	345.5	27.6	847	A53800	protein kinase hom
21	344	27.5	390	T01451	protein-tyrosine k
22	341	27.2	410	B35670	hypothetical prote
23	340	27.2	387	T22511	hypothetical prote
24	336	26.8	356	S61766	protein kinase ATN
25	329.5	26.3	357	C84856	probable protein k
26	329.5	26.3	545	T05675	hypothetical prote
27	328.5	26.2	377	T46149	protein kinase ATN
28	328.5	26.2	736	T05137	protein kinase hom
29	324	25.9	982	T06576	probable protein k

ALIGNMENTS

RESULT 1

JC5957
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1c - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 17-Mar-2000
C:Accession: JC5957
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5957
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-567 <SAK>
A:Cross-references: DDBJ:AB009358
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: phosphotransferase

Query Match	100.0%	Score 1252;	DB 2;	Length 567;
Best Local Similarity	100.0%	Pred. No. 1.5e-72;		
Matches	228;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	VELRQLSRVNHNPVVKLYGACLN	PVCLVMEYABGGSLYNVLHGAEPLPYTTAAHMSWCL	60
Db	76	VELRQLSRVNHNPVVKLYGACLN	PVCLVMEYABGGSLYNVLHGAEPLPYTTAAHMSWCL	135
Qy	61	QCSQGVAYLHSMOPKALIH	RDLKPPNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAW	120
Db	136	QCSQGVAYLHSMOPKALIH	RDLKPPNLLVAGGTVLKICDFGTACDIQTHMTNKGSAAW	195
Qy	121	MAPEVFGSNYSKCDVFSWGI	ILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	180
Db	196	MAPEVFGSNYSKCDVFSWGI	ILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	255
Qy	181	PKPIESLMTFCWSKQSPQSR	PSMBEIVKIMTHLMRYFPFGADEPLQYPCQ	228
Db	256	PKPIESLMTFCWSKQSPQSR	PSMBEIVKIMTHLMRYFPFGADEPLQYPCQ	303

RESULT 2

JC5955
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1a - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
C:Accession: JC5955
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243, 545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5955
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-579 <SAK>
A:Cross-references: DDBJ:AB009356; NID:g2924623; PIDN:BA025025.1; PID:g2924624
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 579;
Best Local Similarity 100.0%; Pred. No. 1.5e-72;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 60
DB VELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 135
QY 61 QCSQGVAVLHSMOPKALIHRLDKPNNLLVAGTVLKI CDGTGTCDIQTHTNNKGSAAW 120
DB QCSQGVAVLHSMOPKALIHRLDKPNNLLVAGTVLKI CDGTGTCDIQTHTNNKGSAAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 180
DB MAPEVEGSGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTKRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 181 PKPIESLMTKRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 303
QY 256 PKPIESLMTKRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 3
JC5956
transforming growth factor-beta activated kinase (EC 2.7.-.-) 1b - human
C:Species: Homo sapiens (man)
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
R:Accession: JC5956
R:Sakurai, H.; Shigemori, N.; Hasegawa, K.; Sugita, T.
Biochem. Biophys. Res. Commun. 243:545-549, 1998
A:Title: TGF-beta-activated kinase 1 stimulates NF-kappaB activation by an NF-kappaB-ind
A:Reference number: JC5955; MUID:98153801; PMID:9480845
A:Accession: JC5956
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-606 <SAK>
A:Cross-references: DDBJ:AB009357; NID:g2924625; PIDN:BA025026.1; PID:g2924626
A:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
C:Keywords: phosphotransferase

Query Match 100.0%; Score 1252; DB 2; Length 606;
Best Local Similarity 100.0%; Pred. No. 1.6e-72;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 60
DB VELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 135
QY 61 QCSQGVAVLHSMOPKALIHRLDKPNNLLVAGTVLKI CDGTGTCDIQTHTNNKGSAAW 120
DB QCSQGVAVLHSMOPKALIHRLDKPNNLLVAGTVLKI CDGTGTCDIQTHTNNKGSAAW 195
QY 121 MAPEVEGSGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 180
DB MAPEVEGSGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPARIMAAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTKRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 181 PKPIESLMTKRCWSKDPSPRSMEEIVKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 4
S68178
mixed-lineage protein kinase 2 (EC 2.7.1.-.-) - human
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 11-Jun-1999
C:Accession: S68178; I38044; S32468
R:Dorow, D.S.; Devereux, L.; Th, G.F.; Price, G.; Nicholl, J.K.; Sutherland, G.R.; Simps
Eur. J. Biochem. 234: 492-500, 1995

A:Title: Complete nucleotide sequence, expression, and chromosomal localisation of huma
A:Reference number: S68178; MUID:96128179; PMID:8536694
A:Accession: S68178
A:Molecule type: mRNA
A:Residues: 1-954 <DOR>
A:Cross-references: EMBL:X0846; NID:g9711419; PIDN:CAA62351.1; PID:g9711420
R:Katon, M.; Hirai, M.; Sugimura, T.; Terada, M.
Oncogene 10, 1447-1451, 1995
A:Title: Cloning and characterization of MST, a novel (putative) serine/threonine kinas
A:Reference number: I38044; MUID:95249256; PMID:7731697
A:Accession: I38044
A:Status: translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-461, 'A', 'V', '465-470', 'S', '472-806', 'R', '808-817', 'A', '819-954 <RES>
A:Cross-references: EMBL:Z48615; NID:g758592; PIDN:CAA8531.1; PID:g758593
R:Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A:Title: Identification of a new family of human epithelial protein kinases containing t
A:Reference number: S32467; MUID:93238756; PMID:8477742
A:Accession: S32468
A:Molecule type: mRNA
A:Residues: 244-464, 'AOAGRPQHPALWL' <DD2>
C:Genetics:
A:Gene: GDB:MLK2; GDB:MST
A:Cross-references: GDB:362654; GDB:624810; OMIM:600137
A:Map position: 19q13.1-19q13.2
C:Superfamily: mixed-lineage protein kinase 2; protein kinase homology; SH3 homology
C:Keywords: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein k
F:23-76/Domains: SH3 homology <SH3>
F:96-364/Domains: protein kinase homology <KIN>
F:104-112/Region: protein kinase ATP-binding motif
F:384-405/Region: leucine zipper motif
F:419-440/Region: leucine zipper motif
F:449-463/Region: basic
F:125,145,222,224/Active site: Lys, Glu, Asp, Lys #status predicted

Query Match 31.3%; Score 392.5; DB 1; Length 954;
Best Local Similarity 40.6%; Pred. No. 7.8e-18;
Matches 88; Conservative 35; Mismatches 77; Indels 17; Gaps 6;

QY 2 ELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 59
DB 145 ELRLSRVNHNPVIVKLYGACINPVCVMEYABGSLYVNLHGAEPLPYTTAAHMSMCL 200
QY 60 QCSQGVAVLHSMOPKALIHRLDKPNNLLVAGTVLKI CDGTGTCDIQTHTNNKGSAAW 111
DB 201 VOYARGMNYLHNDAPVPIHRLDKSINILILEAIENHMLADTVLKITDGLAREHMKTK 260
QY 112 TNNKGSAAWMAPEVEGSGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPARIMAAV-HN 170
DB 261 MSAAGTYAMMAPEVEIRLSLFSKSDVWSFGVLWELLGVEFYREI--DALAVAGVAMN 318
QY 171 GTRPPLIKNLPKPIESLMTKRCWSKDPSPRSMEEIVK 207
DB 319 KLTLPISPTCPPEPFARILKECDMDPDRHPRDGSILK 355

RESULT 5
G84635
Probable protein kinase (imported) - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C:Accession: G84635
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanhaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84635
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-407 <STO>

A;Cross-references: GB:AE002093; NID:g4337195; PIDN:AAD18109.1; GSPDB:GN00139
C;Genetics:
A;Gene: At2g24360
A;Map position: 2
C;Superfamily: kinase-related transforming protein; protein kinase homology

Query Match 29.9%; Score 374; DB 2; Length 407;
Best Local Similarity 39.4%; Pred. No. 5.6e-17;
Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHPIVVKLYGACLNPNV--CLWMEYAEAGGSLYNYL-----HGAEPLPYTTAAHA 55
Db 175 EVSMLANLKHPIVRFIGACRKPMMWCIVTEYAKGSVRQFLTRQNRVPLKL-----A 229

QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFGTA-CDIQTH-WTN 113
Db 230 VKQALDVARGMAYVHG---RNFIHRDLKSDNLLISADKSI-KIADFGVARIETQEGMTP 285

QY 114 NKGSAAMWAPVFEFGSNYSKCDVFSWGIIWEVITRRKPFDEIGG--PAFRIMWAVHG 171
Db 286 ETGTYRWMAPEMIQHRAYNQKVDVYSGVILWELITGLLPQNTAVQAFAV---VNRG 342

QY 172 TRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIM 209
Db 343 VRPTVPNDCLPVLSDIMTRCWDANPEVRPCFVGVVKLL 380

RESULT 6
T10671
protein kinase homolog F6E21.90 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 19-May-2000
C;Accession: T10671
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft
submitted to the Protein Sequence Database, June 1999
A;Reference number: Z16533
A;Accession: T10671
A;Molecule type: DNA
A;Residues: 1-412 <BV>
A;Cross-references: EMBL:AL049914; GSPDB:GN00062; ATSP:F6E21.90
A;Experimental source: cultivar Columbia; BAC clone F6E21
C;Genetics:
A;Gene: ATSP:F6E21.90
A;Map position: 4
A;Introns: 300/2
C;Superfamily: kinase-related transforming protein; protein kinase homology
F:135-392/Domain: protein kinase homology <KIN>

Query Match 29.8%; Score 373; DB 2; Length 412;
Best Local Similarity 39.9%; Pred. No. 6.6e-17;
Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRQLSRVNHPIVVKLYGACLNPNV--CLWMEYAEAGGSLYNYL-----HGAEPLPYTTAAHA 55
Db 180 EVSMLAFLKHPNIVRFIGACIKPMWCIVTEYAKGSVRQFLTKRQNRVPLKL-----A 234

QY 56 MSWCLOCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFGTA-CDIQTH-WTN 113
Db 235 VQQALDVARGMAYVHE---RNFIHRDLKSDNLLISADRSI-KIADFGVARIETQEGMTP 290

QY 114 NKGSAAMWAPVFEFGSNYSKCDVFSWGIIWEVITRRKPFDEIGG--PAFRIMWAVHG 171
Db 291 ETGTYRWMAPEMIQHRPYTKQVDVYSGVILWELITGLLPQNTAVQAFAV---VNRG 347

QY 172 TRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIM 209
Db 348 VRPTVPADCLPVLGEIMTRCWDADPEVRPCFABIVNLL 385

RESULT 7
JU0229
mixed-lineage protein kinase 1 - human
C;Species: Homo sapiens (man)

C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 17-Mar-1999
A;Accession: S32467; JU0229
R;Dorow, D.S.; Devereux, L.; Dietzsch, E.; de Kretser, T.
Eur. J. Biochem. 213, 701-710, 1993
A;Title: Identification of a new family of human epithelial protein kinases containing tv
A;Reference number: S32467; MUID:93238756; PMID:8477742
A;Accession: S32467
A;Molecule type: mRNA
A;Residues: 1-394 <DO2>
C;Genetics:
A;Gene: GDB:MLK1
A;Cross-references: GDB:141921; OMIM:600136
A;Map position: 14q24.3-14q31
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C;Key words: ATP; leucine zipper; phosphotransferase; serine/threonine-specific protein K;
F:1-269/Domain: protein kinase homology <KIN>
F:1-268/Domain: catalytic <CAT>
F:9-17/Region: protein kinase ATP-binding motif
F:289-310/Region: leucine zipper motif
F:324-345/Region: leucine zipper motif
F:354-368/Region: basic

Query Match 29.0%; Score 363.5; DB 2; Length 394;
Best Local Similarity 37.7%; Pred. No. 2.5e-16;
Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

QY 2 ELRQLSRVNHPIVVKLYGACLNPNV--VCLWMEYAEAGGSLYNYLHGAEPLPYTTAAHAMSWC 59
Db 50 EAKLFAMLKHPNIIALRGVCLKEPNLCVMEFARGGPLNRLSGKRIPPD1----LVNWA 105

QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLV-----AGGTVLKICDRGTACD-IQTHM 111
Db 106 VQIARGMNYLHDAIVPIIHRDLKSSNIILOKVENGDLSNLIKILITDFGLAREWHRTTK 165

QY 112 TNKGSAAWMAPEVFEFGSNYSKCDVFSWGIIWEVITRRKPFDEIGGPAFRIMWAV-HN 170
Db 166 MSAAGTYWMAPEVIRASMFSGSDVMSVGVLLWELLTGEVPRGIDG--LRVAYGVAMN 223

QY 171 GTRPPLIKNLKPKIESLMTRCWSKDPSPRSMEEIVKIMT 210
Db 224 KLALPIPTCTPEPAFLMEDCWNPDPHSRPSFTNILDQLT 263

RESULT 8
T18287
protein-tyrosine kinase (EC 2.7.1.112) - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 04-Feb-2000
C;Accession: T18287
R;Adler, K.
submitted to the EMBL Data Library, July 1996
A;Reference number: Z18856
A;Accession: T18287
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-1338 <ADL>
A;Cross-references: EMBL:U64830; NID:g1468982; PID:g1468983; PIDN:AAB04999.1
A;Introns: 1181/3
C;Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase

Query Match 28.9%; Score 362; DB 2; Length 1338;
Best Local Similarity 38.1%; Pred. No. 9e-16;
Matches 80; Conservative 41; Mismatches 73; Indels 16; Gaps 6;

QY 2 ELRQLSRVNHPIVVKLYGACLNPNVCLWMEYAEAGGSLYNYLH---GAEPPLPYTTAAHMS 57
Db 1100 EVSSLKSHHPNVVTFMGARIDPPCIFTEYLQGGSLLDVLIHQIKILNPLMMYKWIHDL- 1158

QY 58 WCLQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFTACDIQTHMT-NNKG 116
Db 1159 -----SLGMEHLHSIQ---MLHRDLTSKNILLDEFKNI-KIADFGLATTLSDDDMTLSGIT 1209

QY 117 SAAMAAEVEFGSNYSSEKCDVFSIGILLMTVEITRRKPFDEIGGAPRIMAAVHNGRRPL 176

Db 1210 NPRWRPELTGGLVYNEKVDVPSGLVVEIYTGKILPFEBLDDTAAAGAAAPEN-YRPAI 1266

QY 177 IKNLKPLESIMTRCWSKDSQRRSMELV 206

Db 1269 PPDCVSLRKILITCKMADSPQRRSFTEIL 1298

RESULT 9
S29851
protein kinase 6 (EC 2.7.1.-) - soybean
C,Species: Glycine max (soybean)
C,Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 24-Sep-1999
C,Accession: S29851; S27760
C,Feng, X.H.; Zhao, Y.; Bottino, P.J.; Kung, S.
Biochim. Biophys. Acta 1172, 200-204, 1993
A,Title: Cloning and characterization of a novel member of protein kinase family from soybean
A,Reference number: S29851; MUID:93176812; PMID:8439562
A,Accession: S29851
A,Molecule type: mRNA
A,Residues: 1-462 <FEN>
A,Cross-references: EMBL:M67449; NID:G170046; PIDN:AAA34002.1; PID:G170047
C,Superfamily: unassigned Ser/thr or Tyr-specific protein kinases; protein kinase homolo
C,Keywords: ATP; phosphotransferase
F,154-419/Domain: protein kinase homology <KIN>

RESULT 10
T46150
protein kinase ATN1-like protein - Arabidopsis thaliana
N/Alternate names: protein T3A5.110
N/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 17-Mar-2000
C/Accession: T46150; T08394
R/Bloeker, H.; Mewes, H.W.; Mayer, K.F.X.; Lemcke, K.; Schueller, C.; Queciet, F.; Salas,
submitted to the Protein Sequence Database, December 1999
A/Accession: T46150
A/Reference number: Z23024
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-570 <BLO>
A/Cross-references: EMBL:ALJ32979
A/Experimental source: cultivar Columbia; BAC clone T3A5
R/Queciet, F.; Rieger, M.; Gabel, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Salanoub
submitted to the Protein Sequence Database, May 1999
A/Reference number: Z16409
A/Accession: T08394
A/Molecule type: DNA
A/Residues: 1-370 <QUE>
A/Cross-references: EMBL:AL049862; GSPDB:GN00061; ATSP:FLIB3.10
A/Experimental source: cultivar Columbia; BAC clone FLIB3
A/Genetics:

A:Gene: ATSP:FL18B3.10
A:Map position: 3
A:Introns: 54/2; 99/3; 154/1; 204/3; 257/3
A:Note: T3A5.110
C:Superfamily: basic fibroblast growth factor receptor 1; immunoglobulin homology; prot

		28.4%;	Score 35.5;	DB 2;	Length 370;
		Best Local Similarity	35.4%;	Pred. No. 7;7e-16;	
Matches	79;	Conservative	46;	Mismatches	81;
				Indels	17;
				Gaps	8;
OY	2	EURQLSRVNHPIYLKYGACLN-P-VCLMEYAEAGSLIVLHGAPLEPYTTAAHMSWCL	60		
Db	85	EYLLSKKHNNIYKFVGACIEPOLIIITVELVEGTLGRFPH-SRPGP-LDLKMSLSPAL	142		
OY	61	OCSCGVALLHSQPKALIRHDKPEPNILLVAGGYLKICDGTGC-DIQHTMTNKKSGAA	119		
Db	143	DISRAEMEVHS---NGITHRDLPNRLLVTGDDLKVAKLADFGILREETRGMTCEAGTSK	199		
OY	120	WMAPVF-----EGSNYSSEKCDVFSWGIIILWEVIITRRKPDEIGPAFRIMAWANG	171		
Db	200	WMAPFVSPEELRVGEKEKEYDHAKDIYSFAVLWLQVLTNEEPFDVPNSLF-VPLVSSQG	258		
OY	172	TRPPLIKULPKRIEISLMTRCWSKSPQRSMEEIYKIMTHLMR	214		
Db	259	KRPILTKT-PDVFPVIVESCMAQPDARPEPKESIVMLTNLLR	300		

```

RESULT 11
JC2363
Protein kinase (EC 2.7.1.37) ZPK - human
N/Alternate names: leucine-zipper protein kinase
C/Species: Homo sapiens (man)
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 02-Feb-2001
C/Accession: JC2363
R/Recomm. U.R.: Pleasure, D.
Biochem. Biophys. Res. Commun. 202, 613-620, 1994
A/Title: Cloning of a novel putative protein kinase having a leucine zipper domain from
A/Reference number: JC2363; MUID:94311945; PMID:8037767
A/Accession: JC2363
A/Molecule type: mRNA
A/Residues: 1-668 <RED>
A/Cross-references: EMBL:U07358
A/Experimental source: Brain
A/Note: The nucleotide sequence for this amino acid sequence is inconsistent with that
of the codon ACC for residue 661 as Pro, the codon GAACACCTCCTCCA for residues 664-668 as
C/Comment: This protein belongs to the family of non-receptor kinase.
A/Genetics:
A/Gene: GDB:ZPK
A/Cross-references: GDB:383963; OMIM:600447
A/Map position: 12q13-12q13
C/Superfamily: Unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homol
C/Keywords: ATP; leucine zipper; nucleotide binding; P-loop; phosphotransferase
F/113-139/Domain: protein kinase homology <KIN>
F/443-471/Region: protein kinase ATP-binding motif
F/538-545/Region: leucine zipper motif
F/152/Active site: Lys #status predicted

Query Match      28.2%; Score 353.5; DB 2; Length 668;
Best Local Similarity 34.7%; Pred. No. 1.7e-15;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQLSRVHNPNITVILYGACINP--VCLVMEYAEAGSGILVNVLHGAEPDLPYTAHAHMSWC 59
DB 161 DIKHLRLKRPNIITTFPGVCTQAPCYCLIMEFCAQGLYEVLAAGRNV--TSLSLVDWS 217
QY 60 LQCSQYAYVYHMQPKALLHRDLKPEVNLIVAGGVVLKICDPGTACDIQTHMTNNK--GS 117
DB 218 MGLAGGNVYHLHK--IHRDLKSPN-MLITYDVKIDFSGKSELSDKSTKMSFAGT 273
QY 118 AAMVAPEVFEFGSYSEKCDVFSMGIIIMWEVITRRKRPDEIDCGAFRIMAV-HNGTRPPL 176
DB 274 VAMVAPEVIRNEVSEKVDVMSFGVAVLWELTETIRPKVDGSA--IIMGVGSNSLHPV 331

```

F:557-888/Region: glycine-serine-proline rich #status predicted

Query Match 28.1%; Score 351.5; DB 2; Length 888;
Best Local Similarity 34.3%; Pred. No. 2.9e-15;
Matches 74; Conservative 44; Mismatches 81; Indels 17; Gaps 8;

QY 177 IKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHL 212
Db 332 PSSCPDGFKILLRQCWNKSPRNPSPFRQ---ILLHL 364

RESULT 12

A55318
serine/threonine protein kinase (EC 2.7.1.-) Dlk - mouse
C:Species: Mus musculus (house mouse)
C>Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 24-Sep-1999
C:Accession: A55318
R:Holzman, L.B.; Merritt, S.E.; Fan, G.
J. Biol. Chem. 269, 30808-30817, 1994
A:Title: Identification, molecular cloning, and characterization of dual leucine zipper
s.
A:Reference number: A55318; MUID:95074107; PMID:7983011
A:Accession: A55318
A:Status: preliminary
A:Residues: 1-888 <HOL>
A:Cross-references: GB:U14636; NID:G602677; PIDN:AAAS7280.1; PID:G602678
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; leucine zipper; phosphotransferase
F:156-404/Domain: protein kinase homology <KIN>
F:164-172/Region: protein kinase ATP-binding motif

Query Match 28.2%; Score 353.5; DB 2; Length 888;
Best Local Similarity 34.7%; Pred. No. 2.2e-15;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQLSRVNHNPINVLKYGACLNLP--VCLVMEYAEAGGSLYVNLHGAEPPLPYTAAHAMSWC 59
Db 194 DIKHLRKLKHPNIITFGVCTQAPCVCIIMEFCAOQQLYEVLRAGRPV---TPSLLDVMS 250

QY 60 LQCSQGVAVLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GS 117
Db 251 MGIAGGMNVLHLK---IHRDLKSPN-MLITYDDVKISDFGTSKELSDKSTKMSFAGT 306

QY 118 AAWMAPEVFEAGSYSEKCDVFWGIIWVITRRKPFDEIGGPAFRIMWAV-HNGTRPPL 176
Db 307 VAWMAPEVIRNEPVSEKVDIWSFGVVLWELLTGEIPYKDVDSA--IIVGVGSNSLHLPV 364

QY 177 IKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHL 212
Db 365 PSSCPDGFKILLRQCWNKSPRNPSPFRQ---ILLHL 397

RESULT 13

JCS399
dual leucine zipper kinase (EC 2.7.1.-) - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 07-Jul-1997 #sequence_revision 18-Jul-1997 #text_change 28-May-1999
C:Accession: JCS399
R:Matsu, N.; Sarkar, G.; Shuto, T.; Marrs, J.; Bronk, J.T.; Mizuno, K.; Bolander, M.E.
Biochem. Biophys. Res. Commun. 229, 571-576, 1996
A:Title: Identification of a dual leucine zipper kinase involved in rat fracture repair.
A:Reference number: JCS399; MUID:97127443; PMID:8954939
A:Accession: JCS399
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-888 <MAT>
C:Comment: This enzyme is involved in regulating cell function in the musculoskeletal system
C:Genetics:
A:Gene: rDlk
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog
C:Keywords: ATP; phosphotransferase
F:54-75/Region: glycine-rich
F:156-454/Domain: kinase catalytic #status predicted <CAT>
F:164-172/Region: protein kinase homology <KIN>
F:421-449/Region: leucine zipper motif
F:472-500/Region: leucine zipper motif

F:557-888/Region: glycine-serine-proline rich #status predicted

Query Match 28.1%; Score 351.5; DB 2; Length 888;
Best Local Similarity 34.3%; Pred. No. 2.9e-15;
Matches 74; Conservative 44; Mismatches 81; Indels 17; Gaps 8;

QY 2 ELRQLSRVNHNPINVLKYGACLNLP--VCLVMEYAEAGGSLYVNLHGAEPPLPYTAAHAMSWC 59
Db 194 DIKHLRKLKHPNIITFGVCTQAPCVCIIMEFCAOQQLYEVLRAGRPV---TPSLLDVMS 250

QY 60 LQCSQGVAVLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK--GS 117
Db 251 MGIAGGMNVLHLK---IHRDLKSPN-MLITYDDVKISDFGTSKELSDKSTKMSFAGT 306

QY 118 AAWMAPEVFEAGSYSEKCDVFWGIIWVITRRKPFDEIGGPAFRIMWAV-HNGTRPPL 176
Db 307 VAWMAPEVIRNEPVSEKVDIWSFGVVLWELLTGEIPYKDVDSA--IIVGVGSNSLHLPV 364

QY 177 IKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHL 212
Db 365 PSSCPDGFKILLRQCWNKSPRNPSPFRQ---ILLHL 397

RESULT 14

D84555
probable protein kinase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: D84555
M:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Frazer, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: D84555
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-546 <STO>
A:Cross-references: GB:AE002093; NID:G6598802; PIDN:AAB80785.2; GSPDB:GN00139
C:Genetics:
A:Gene: At2g17700
A:Map position: 2

Query Match 28.0%; Score 350; DB 2; Length 546;
Best Local Similarity 36.3%; Pred. No. 2.4e-15;
Matches 81; Conservative 41; Mismatches 83; Indels 18; Gaps 8;

QY 3 LRQLS-----RVNHPNIVKLYGACLN--PVCLVMEYAEAGGSLYVNLHGAEPPLPYTAA 53
Db 325 LRFPSEQVFIMRKVRHKNNVQFLGACTRSPCLTCTVTEFMARGSIYDFLHKQKCA--FKLQ 382

QY 54 HAMSVCQLCSQGVAVLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTA-CDIOTH-M 111
Db 383 TLLKVALDVAKGMSYLHQ---NNIIHRDLKTANLLMDEHGLV-KVADFGVARVQIESGVM 438

QY 112 TNNKGSAAWMAPEVFEAGSYSEKCDVFWGIIWVITRRKPFDEIGGPAFRIMWAVHNG 171
Db 439 TAETGTYRWMAPEVIEHKPNYHKADVFSYAIVLWELLTGDIPYAFV-TPLOAAVGVVQKG 497

QY 172 TRPPLIKNLKPPIESLMTRCWSKDPQSRPSMEEIVKIMTHLMR 214
Db 498 LRPKTPKTPHKVKGLLERCHQDPEORPLPEEIIEMLQOIMK 540

RESULT 15

T48115
protein kinase ATRMK1 (EC 2.7.1.-) [imported] - Arabidopsis thaliana
N:Alternate names: protein F16M2.110
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2001
C:Accession: T48115; T51942
R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemcke, K.

submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24459

A:Accession: T48115
A:Molecule type: DNA

A:Residues: 1-391 <RIE>

A:Cross-references: EMBL:AL18648; NID:g7523397; PIDN:CA86427.1; PID:g7523408
A:Experimental source: cultivar Columbia; BAC clone F16M2

R:Richimura, K.; Mizoguchi, T.; Shinozaki, K.

Plant Sci. 130, 171-179, 1997

A:Title: AtMKI, an Arabidopsis protein kinase related to mammal mixed-lineage kinases a

A:Reference number: Z24427

A:Accession: T51942

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-391 <ICH>

A:Cross-references: EMBL:AB006810; NID:g2351096; PIDN:BA22079.1; PID:g2351097

A:Experimental source: cultivar Columbia

C:Genetics:

A:Map position: 3

A:Introns: 109/1; 149/3; 220/2; 278/3; 323/3

A>Note: F16M2.110

C:Superfamily: Kinase-related transforming protein; protein kinase homology

C:Keywords: phosphotransferase

Query Match 27.8%; Score 348; DB 2; Length 391;
Best Local Similarity 33.7%; Pred. No. 2.4e-15;
Matches 82; Conservative 35; Mismatches 76; Indels 50; Gaps 8;

QY 2 ELRQLSRVNHPIVKLYGACL-----NPVCLVMEYAGGSly 38

Db 135 EVAVMQKLDHPNVTKFIGASMGTSDLRIPPAGDTGGRGNAGHAPARACCVVEYVAGGTLK 194

QY 39 NVLHGAELPYTA-----AHAMSWCLQCSQGVAYLHSMQPKALIRDLKPNLLIVAGT 94

Db 195 KFL-----IKKYRAKLPIKDVITQALDLARGSLYLS--KAIVHRDVXSENMLLOPNKT 246

QY 95 VLKICDPFTA---CDIQHTMTNNKSAAMAPEVEFGSNYSSEKCDVPSFGITLMEVITRR 151

Db 247 -LKINADFGVARVEAQNPQDMTGETGTGLGMAPEVLEGGPKYRKCDVYSFGVCLMEIYICD 305

QY 152 KP-----FDEIGGPAFRIMMAVHNGTRPPLIKNLKPKIESLMTRCWXPQSRPSMERIV 206

Db 306 MPYADCSFAEISHA-----VVHRNLRPEIPKCCPHAVANIMKRCWDPNDRPEMEEVV 359

QY 207 KIM 209

Db 360 KLL 362

Search completed: December 10, 2002, 03:58:57
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:50:45 ; Search time 16 Seconds
(without alignments)
591.038 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELQLSRVNHPIVKLYGA.....MTHLMRYPPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1252	100.0	579	1 M3K7_MOUSE	Q62073 mus musculus
2	1252	100.0	606	1 M3K7_HUMAN	Q43318 homo sapien
3	435.5	34.8	393	1 M3K7_DROME	P83104 drosophila
4	392.5	31.3	954	1 M3KA_HUMAN	Q02779 homo sapien
5	363.5	29.0	394	1 M3K9_HUMAN	P80192 homo sapien
6	355.5	28.4	888	1 M3K9_RAT	Q63796 rattus norv
7	353.5	28.2	859	1 M3KC_HUMAN	Q12852 homo sapien
8	353.5	28.2	888	1 M3KC_MOUSE	Q60700 mus musculus
9	341	27.2	410	1 KYK2_DICDI	P18161 dictyosteli
10	323	25.8	821	1 CTR1_ARATH	Q05609 arabidopsis
11	307.5	24.6	1115	1 RET_MOUSE	P35546 mus musculus
12	304.5	24.3	1114	1 RET_HUMAN	P07949 homo sapien
13	301.5	24.1	630	1 TEC_MOUSE	P24604 mus musculus
14	301.5	24.1	631	1 TEC_HUMAN	P42680 homo sapien
15	299.5	23.9	625	1 ITK_MOUSE	Q03526 mus musculus
16	298.5	23.8	620	1 ITK_HUMAN	Q08881 homo sapien
17	294.5	23.5	2347	1 KROS_HUMAN	P08922 homo sapien
18	292.5	23.4	402	1 KROS_AVISU	P00529 avian sarco
19	292	23.3	507	1 KROS_CHICK	P08941 gallus gall
20	290	23.2	1584	1 KYK1_DICDI	P18160 dictyosteli
21	288.5	23.0	628	1 TESK_RAT	Q63572 rattus norv
22	288	23.0	1307	1 PHYI_CERPU	P25948 ceratodon p
23	286	22.8	804	1 FPS_DROME	P18106 drosophila
24	284.5	22.7	832	1 ANR3_HUMAN	P57078 homo sapien
25	283.5	22.6	527	1 TXK_MOUSE	P42882 mus musculus
26	280	22.4	822	1 FGRI_RAT	Q04589 rattus norv
27	279	22.3	467	1 MATK_RAT	P41243 rattus norv
28	278.5	22.2	590	1 BTKL_DROME	P08630 drosophila
29	278	22.2	822	1 FGRI_HUMAN	P11362 homo sapien
30	278	22.2	822	1 FGRI_MOUSE	P16092 mus musculus
31	275	22.0	984	1 EPA3_RAT	Q08680 rattus norv
32	274.5	21.9	801	1 FGRI_MOUSE	Q61851 mus musculus
33	271.5	21.7	1062	1 CC7_SCHPO	P41892 schizosacch

34	271	21.6	604	1 KRAA_MOUSE	P04627 mus musculus
35	271	21.6	604	1 KRAA_RAT	P14056 rattus norv
36	270.5	21.6	503	1 HCK_RAT	P50545 rattus norv
37	270	21.6	806	1 FGR3_HUMAN	P22607 homo sapien
38	269.5	21.5	507	1 MATK_HUMAN	P42679 homo sapien
39	269.5	21.5	626	1 TESK_HUMAN	Q15569 homo sapien
40	269	21.5	533	1 FPS_AVISP	P00541 avian sarco
41	269	21.5	819	1 FGRI_CHICK	P21804 gallus gall
42	269	21.5	983	1 EPA3_HUMAN	P29320 homo sapien
43	268	21.4	830	1 PKN2_MYXXA	P54736 myxococcus
44	268	21.4	1052	1 FAK1_HUMAN	Q05397 homo sapien
45	268	21.4	1052	1 FAK1_MOUSE	P34152 mus musculus

ALIGNMENTS

```
RESULT 1
M3K7_MOUSE
ID M3K7_MOUSE STANDARD; PRT; 579 AA.
AC Q62073;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Mitogen-activated protein kinase kinase kinase 7 (EC 2.7.1.-)
DE (transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAP3K7 OR TAK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96123277; PubMed=8533096;
RA Yamaguchi K., Shirakabe K., Shibuya H., Irie K., Ohishi I., Ueno N.,
RA Taniguchi T., Nishida E., Matsumoto K.;
RT "Identification of a member of the MAPKKK family as a potential
RL mediator of TGF-beta signal transduction.";
RL Science 270:2008-2011(1995).
CC -!- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MAPKKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D76446; BA01184.1; -
CC HSSP; P08631; 1AD5.
CC MGD; MGI:1346877; Map3k7.
CC InterPro; IPR000719; Euk_pkinase.
CC InterPro; IPR004040; STY_pkinase.
CC InterPro; IPR002290; Ser_thr_pkinase.
CC Pfam; PF00069; pkinase; 1.
CC ProDom; PD000001; Euk_pkinase; 1.
CC SMART; SM00221; STYKc; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
CC PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
CC Transfaser; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 8 16 POLY-SER.
FT DOMAIN 36 291
FT NP_BIND 42 50 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
```

```

SQ SEQUENCE 579 AA, 64227 MW, 97C8F6F3C8E283EE CRC64;
Query Match 100.0%; Score 1252; DB 1; Length 579;
Best Local Similarity 100.0%; Pred. No. 3.7e-111;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVLYGACLNPCVLMVEYAGSLVNVHLGAPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHNPVIVLYGACLNPCVLMVEYAGSLVNVHLGAPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIRHDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIRHDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAM 195
QY 121 MAPEVEGNSYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPARIMAAVANGTRPPLIKNL 180
DB 196 MAPEVEGNSYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPARIMAAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPRSMEEIVKIMTHLMRFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPRSMEEIVKIMTHLMRFPGADEPLQYPCQ 303

RESULT 2
M3K7 HUMAN
ID M3K7 HUMAN STANDARD; PRT; 606 AA.
AC 04318; 04317; 04319;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 7 (EC 2.7.1.-)
DE (Transforming growth factor-beta-activated kinase 1) (TGF-beta-
DE activated kinase 1).
GN MAPK7 OR TAK1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]

SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RX MEDLINE=98153801; PubMed=9480845;
RA Sakurai H., Shigemori N., Hasegawa K., Sugita T.;
RT "TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
RT NF-kappa B-inducing kinase-independent mechanism.";
RL Biochem. Biophys. Res. Commun. 243:545-549(1998).
CC -1- FUNCTION: CAN PHOSPHORYLATE AND ACTIVATE YET UNDEFINED MARKKS.
CC MEDIATOR OF TGF-BETA SIGNAL TRANSDUCTION. STIMULATES NF-KAPPA B
CC ACTIVATION.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 1A, 1B (SHOWN HERE) AND 1C; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -1- SIMILARITY: IN THE C-TERMINAL SECTION; STRONG, TO C21ORF7.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB009357; BAA25026.1; -
DR EMBL; AB009356; BAA25025.1; -
DR EMBL; AB009358; BAA25027.2; -
DR HSSP; P08631; 1AD5.
DR HGNC; 6859; MAP3K7.
DR MIM; 602614; -
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STR_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; kinase; 1.
```

```

DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferrase, Serine/threonine-protein kinase; ATP-binding;
KW Alternative splicing.
FT DOMAIN 8 14 POLY-SER.
FT NP_BIND 36 291 PROTEIN KINASE.
FT NP_BIND 42 50 ATP (BY SIMILARITY).
FT BINDING 63 63 ATP (BY SIMILARITY).
FT ACT_SITE 156 156 BY SIMILARITY.
FT VARSPIC 404 430 MISSING (IN ISOFORM 1A).
FT VARSPIC 509 518 PLAPCNSKE -> ARTSCRTSPG (IN ISOFORM 1C).
FT VARSPIC 519 606 MISSING (IN ISOFORM 1C).
SQ SEQUENCE 606 AA; 67196 MW; 3DBF8147CD174013 CRC64;

Query Match 100.0%; Score 1252; DB 1; Length 606;
Best Local Similarity 100.0%; Pred. No. 3.9e-111;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRQLSRVNHNPVIVLYGACLNPCVLMVEYAGSLVNVHLGAPLPYTTAAHMSWCL 60
DB 76 VELRQLSRVNHNPVIVLYGACLNPCVLMVEYAGSLVNVHLGAPLPYTTAAHMSWCL 135
QY 61 QCSQGVAYLHSMQPKALIRHDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAM 120
DB 136 QCSQGVAYLHSMQPKALIRHDLKPPNLLVAGTVLKICDFGTACDIQTHMTNNGSAAM 195
QY 121 MAPEVEGNSYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPARIMAAVANGTRPPLIKNL 180
DB 196 MAPEVEGNSYSKCDVFSWGIIIMVEVITRRKPFDEIGGAPARIMAAVANGTRPPLIKNL 255
QY 181 PKPIESLMTRCWSKDPSPRSMEEIVKIMTHLMRFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCWSKDPSPRSMEEIVKIMTHLMRFPGADEPLQYPCQ 303

RESULT 3
M3K7 DROME
ID M3K7 DROME STANDARD; PRT; 393 AA.
AC P83104;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative mitogen-activated protein kinase kinase 7
DE (EC 2.7.1.-).
GN TAK1L.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]

SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adamatsides P.G., Scherer S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
```


DB 274 VAMMAPEVIRNEPSEKVIDWSPGVLMWELLTGEIPYKVDSSA--IIMGVGSNSLHLPV 331
QY 177 IKNLPRKPIESLMTRCWSKDPSPRSMEEIYKIMTHL 212
DB 332 PSSCPDGFKILLRQCWNKSPRRNRPSPFRQ---ILLHL 364

RESULT 8
MKC_MOUSE
ID MKC_MOUSE STANDARD; PRT; 888 AA.
AC Q60700; P70286;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Mitogen-activated protein kinase kinase 12 (EC 2.7.1.37)
DE (Leucine-zipper protein kinase) (ZPK) (Dual leucine zipper bearing kinase) (DLK).
GN MAP3K12 OR ZPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CD-1; TISSUE=Brain;
RX MEDLINE=95074107; PubMed=7983011;
RA Holman L.B., Merritt S.E., Fan G.;
RT "Identification, molecular cloning, and characterization of dual leucine zipper bearing kinase. A novel serine/threonine protein kinase that defines a second subfamily of mixed lineage kinases.";
RL J. Biol. Chem. 269:30808-30817(1994);
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X Swiss Webster; TISSUE=Brain;
RX MEDLINE=96365388; PubMed=8769565;
RA Blouin R., Beaudoin J., Bergeron P., Nadeau A., Grondin G.;
RT "Cell-specific expression of the ZPK gene in adult mouse tissues.";
RL DNA Cell Biol. 15:631-642(1996).
RN [3]
RP PHOSPHORYLATION, AND MUTAGENESIS OF LYS-185 AND GLU-192.
RX MEDLINE=96279269; PubMed=8663324;
RA Mata M., Merritt S.E., Fan G., Yu G.G., Holman L.B.;
RT "Characterization of dual leucine zipper-bearing kinase, a mixed lineage kinase present in synaptic terminals whose phosphorylation state is regulated by membrane depolarization via calcineurin.";
RL J. Biol. Chem. 271:16888-16896(1996).
CC -I- FUNCTION: May be an activator of the JNK/SAPK pathway.
CC Phosphorylates beta-casein, histone 1 and myelin basic protein in vitro.
CC -I- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -I- COFACTOR: Magnesium.
CC -I- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
CC -I- TISSUE SPECIFICITY: Expressed in brain, kidney, lung, heart, testis, gastrointestinal tract, stomach, liver and pancreas.
CC Within the nervous system, predominantly expressed in neurons and enriched in synaptic terminals.
CC -I- PTM: Autophosphorylated on Ser/Thr. Phosphorylated in cytosol under basal conditions and dephosphorylated when membrane-associated.
CC -I- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U14636; AAA57280.1; -
CC EMBL; U23789; AAB17123.1; -
CC HSRP; P12931; 1PMK.

DR MGD; MG1.1346881; Map3k12.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP, FALSE_NEG.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS50011; PROTEIN KINASE DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Magnesium; Membrane.
FT DOMAIN 158 399
FT NP_BIND 164 172
FT BINDING 185 185
FT ACT_SITE 269 269
FT DOMAIN 56 62
FT DOMAIN 668 671
FT DOMAIN 698 701
FT DOMAIN 753 758
FT MOTIFAGEN 185 185
FT MOTIFAGEN 192 192
FT CONFLICT 18 18
FT CONFLICT 28 29
FT CONFLICT 382 382
FT CONFLICT 494 495
FT CONFLICT 517 517
FT CONFLICT 794 794
SQ SEQUENCE 888 AA; 96083 MW; CFCECFID34F899ABB CRC64;
Query Match 28.2%; Score 353.5; DB 1; Length 888;
Best Local Similarity 34.7%; Pred. No. 9,7e-26;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;
QY 2 ELRQSRVNPVNVKLYGACLP--VCLWMEYAEGLYNNVHGAPLPYTAAHAMSYC 59
DB 194 DIKRLKRLKPNITTRKGVGTQAPCYCIIMFPCAQOLIEVLRAGPV--TPSLLVDS 250
QY 60 LQCSQGVAVLHSMQPKALIHRLDKPNNLLVAGTVLKI CDFGTACDIQTHMTNNK- 117
DB 251 MGJAGGMNVYLHKK--IIHRDLKSPN-MLITVDVVKISDFQTSKELSDKSTKMSFACT 306
QY 118 AAMMAPEVFEFGSNYSKCVFSGVILMEVITRKRFDPDGGAPFIMAV-HNGRRPL 176
DB 307 VAMMAPEVIRNEPSEKVIDWSPGVLMWELLTGEIPYKVDSSA--IIMGVGSNSLHLPV 364
QY 177 IKNLPRKPIESLMTRCWSKDPSPRSMEEIYKIMTHL 212
DB 365 PSSCPDGFKILLRQCWNKSPRRNRPSPFRQ---ILLHL 397

RESULT 9
KYK2_DICDI
ID KYK2_DICDI STANDARD; PRT; 410 AA.
AC P181E1;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase 2 (EC 2.7.1.112) (Fragment).
GN PYKB OR DPYK2.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in Dictyostelium discoideum.";
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -I- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
CC -I- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -----

```

CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M33784; AAA33203.1; --
CC PIR: B35670; B35670.
CC HSSP: P08631; IAD5.
CC DICTYDB: DD03011; pykB.
CC InterPro: IPR000719; Euk_pkinase.
CC InterPro: IPR004040; STY_pkinase.
CC InterPro: IPR001245; Tyr_pkinase.
CC Pfam: PF00069; pkinase; 1.
CC ProDom: PD000001; Euk_pkinase; 1.
CC SMART: SM00221; STYK; 1.
CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
CC Transferrase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
KW NON_TER
FT DOMAIN 108 381 PROTEIN_KINASE.
FT NP_BIND 114 122 ATP (BY SIMILARITY).
FT BINDING 135 135 ATP (BY SIMILARITY).
FT ACT_SITE 232 232 BY SIMILARITY.
SQ SEQUENCE 410 AA; 46386 MW; E9391BB605B9AEC1 CRC64;

Query Match 27.2%; Score 341; DB 1; Length 410;
Best Local Similarity 34.6%; Pred. No. 6.1e-25;
Matches 74; Conservative 42; Mismatches 84; Indels 14; Gaps 7;

QY 2 ELRLQSRVHPNIVKLYGACLNPCVLVMEYAEAGSGSYNVLH---GAEPLPYTAHAHMSW 58
DB 153 EVQNLKKNHQNIVMPGICAYKCAPCITETVMAGSGSYNLHNSPTPKYKYSFPLVLKM 212
QY 59 CLOCSGQVAYLHSMQPKALIHRLDKPPNLLVAGGVVLKICDPGTACD--IOTHMTNNGK 116
DB 213 ATDMALGLHLHSHI---TIVHRLDTQNLILDELGNI-KISDFGLSAEKSRGSGMTMNG 268
QY 117 ---SAAMAPVEPEG-SNYSEKCDVFSWGIILWEVITRRKPFDEIGGPAFRIMVAHNGT 172
DB 269 GICNPRWRPELTKNLGHYSEKVDVYCFSLVWVEIITGPIPSDLGD-SQSAQVAYAGL 327
QY 173 RPPLIKNLKPKIESLMTWCWKPDSQRPMSWEIV 206
DB 328 RPPPEYCDPELKLTLTQCWEADPNDRPPPTIV 361

RESULT 10
CTR1_ARATH STANDARD; PRT; 821 AA.
AC Q05609;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Serine/threonine-protein kinase CTR1 (EC 2.7.1.37).
GN CTR1 OR AT5G03730 OR F17C15.150.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
ON [1]
OR SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia; TISSUE=Seedling;
RX MEDLINE=93161417; PubMed=8431946;
RA Kieber J.J., Rothenberg M., Roman G., Feldmann K.A., Ecker J.R.;
RA "CTR1, a negative regulator of the ethylene response pathway in
RT Arabidopsis, encodes a member of the raf family of protein kinases.";
RL Cell 72:427-441(1993).

```

```

RN RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016721; PubMed=11130714;
RA Tabata S., Kaneko T., Nakamura Y., Kotani H., Kato T., Asamizu E.,
RA Miyajima N., Sasamoto S., Kimura T., Hosouchi T., Kawashima K.,
RA Kohara M., Matsumoto M., Matsumo A., Muraki A., Nakayama S.,
RA Nakazaki N., Naruo K., Okumura S., Shingo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Sato S., de la Bastide M.,
RA Huang E., Spiegel L., Gnoj L., O'Shaughnessy A., Preston R.,
RA Habermann K., Murray J., Johnson D., Rohlfing T., Nelson J.,
RA Stoneking T., Pepin K., Spieth J., Sekhon M., Armstrong J., Becker M.,
RA Belter E., Cordum H., Cordes M., Courtney L., Courtney W., Dante M.,
RA Du H., Edwards J., Fryman J., Haakensen B., Lamar E., Latreille P.,
RA Leonard S., Meyer R., Mulvaney E., Ozersky P., Riley A., Strommatt C.,
RA Wagner-McPherson C., Wollam A., Yoakum M., Bell M., Dedhia N.,
RA Parnell L., Shah R., Rodriguez M., Hoon See L., Vil D., Baker J.,
RA Kirchoff K., Toth K., King L., Bahret A., Miller B., Marra M.,
RA Martienssen R., McCombie W.R., Wilson R.K., Murphy G., Bancroft I.,
RA Volckaert G., Wambutt R., Duesterhoeft A., Stiekema W., Pohl T.,
RA Entian K.-D., Terry N., Hartley N., Bent E., Johnson S.,
RA Langham S.-A., McCullagh B., Robben J., Grymonprez B., Zimmermann W.,
RA Ransperger U., Wedler H., Balke K., Wedler E., Peters S.,
RA van Staveren M., Dirkee W., Moolijman P., Klein Lankhorst R.,
RA Weitznegger T., Bothe G., Rose M., Hauf J., Berner S., Hempel S.,
RA Feldpausch M., Lamberth S., Villarroel R., Gielen J., Ardiles W.,
RA Bents O., Lemcke K., Kolesov G., Mayer K.F.X., Rudd S., Schoof H.,
RA Schueller C., Zaccaria P., Mewes H.-W., Bevan M., Fransz P.F.;
RT "Sequence and analysis of chromosome 5 of the plant Arabidopsis
thaliana.";
RL Nature 408:823-826(2000).
CC -!- FUNCTION: ACTS AS A NEGATIVE REGULATOR IN THE ETHYLENE RESPONSE
PATHWAY.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH SEEDLINGS AND ADULT PLANTS.
CC -!- MISCELLANEOUS: CTR1 MUTANTS DISPLAY ETHYLENE-TREATED PHENOTYPES,
RESULTING IN PLANTS WITH SMALL, UNEXPANDED LEAVES AND WHOSE SEED
COTYLEDON GROWTH IS IMPAIRED.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
MLT/RAF SUBFAMILY.
-----
This SWISS-PROT entry is copyrighted. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL: L08789; AAA32779.1; --
EMBL: L08790; AAA32780.1; --
EMBL: AL162506; CAB82938.1; --
InterPro: IPR000719; Euk_pkinase.
InterPro: IPR004040; STY_pkinase.
Pfam: PF00069; pkinase; 1.
ProDom: PD000001; Euk_pkinase; 1.
SMART: SM00221; STYK; 1.
PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
Transferrase; Serine/threonine-protein kinase; ATP-binding.
KW DOMAIN 65 69 POLY-GLY.
FT DOMAIN 135 141 POLY-GLY.
FT DOMAIN 551 809 PROTEIN_KINASE.
FT NP_BIND 557 565 ATP (BY SIMILARITY).
FT BINDING 578 578 ATP (BY SIMILARITY).
FT ACT_SITE 676 676 BY SIMILARITY.
FT MUTAGEN 596 596 E->K: IN CTR1-4; EXHIBITS ETHYLENE-
TREATED PHENOTYPE.
FT MUTAGEN 694 694 D->E: IN CTR1-1; EXHIBITS ETHYLENE-
TREATED PHENOTYPE.
SQ SEQUENCE 821 AA; 90306 MW; 2922D3DCD0CC15BC CRC64;

```


RP SEQUENCE OF 1-280 FROM N.A.
RX MEDLINE=89282215; PubMed=2660074;
RA Takahashi M.;
RT "Isolation of ret proto-oncogene cDNA with an amino-terminal signal
sequence.";
RL Oncogene 4:805-806(1989).
RN [2]
RP SEQUENCE OF 255-1114 FROM N.A.
RX MEDLINE=90272230; PubMed=3078962;
RA Takahashi M., Buma Y., Iwamoto T., Inaguma Y., Ikeda H., Hiai H.;
RT "Cloning and expression of the ret proto-oncogene encoding a tyrosine
kinase with two potential transmembrane domains.";
RL Oncogene 3:571-578(1988).
RN [3]
RP SEQUENCE OF 588-1063 FROM N.A.
RX MEDLINE=87257826; PubMed=3037315;
RA Takahashi M., Cooper G.M.;
RT "ret transforming gene encodes a fusion protein homologous to
tyrosine kinases.";
RL Mol. Cell. Biol. 7:1378-1385(1987).
RN [4]
RP REVIEW ON HSCR VARIANTS.
RX MEDLINE=98023959; PubMed=9359036;
RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;
RT "Mutations in Hirschsprung disease: when does a mutation contribute to
the phenotype.";
RL Eur. J. Hum. Genet. 5:180-185(1997).
RN [5]
RP REVIEW ON VARIANTS.
RX MEDLINE=97220587; PubMed=9067749;
RA Eng C., Mulligan L.M.;
RT "Mutations of the RET proto-oncogene in the multiple endocrine
neoplasia type 2 syndromes, related sporadic tumours, and
hirschsprung disease.";
RL Hum. Mutat. 9:97-109(1997).
RN [6]
RP VARIANTS MEN2A/MTC W-611; S-618; R-620; Y-620 AND R-634.
RX MEDLINE=93372843; PubMed=8103403;
RA Donis-Keller H., Dou S., Chi D., Carlson K.M., Toshima K.,
RA Laifmore T.C., Howe J.R., Moley J.F., Goodfellow P., Wellis S.A. Jr.;
RT "Mutations in the RET proto-oncogene are associated with MEN 2A and
FMTC.";
RL Hum. Mol. Genet. 2:851-856(1993).
RN [7]
RP VARIANTS G-618; 632-DVR-634; G-634; F-634; Y-634 AND S-634.
RX MEDLINE=93275414; PubMed=8099202;
RA Mulligan L.M., Kwok J.B.J., Healey C.S., Elsdon M.J., Eng C.,
RA Gardner E., Love D.R., Mole S.E., Moore J.K., Papi L., Ponder M.A.,
RA Telenius H., Tunnacliffe A., Ponder B.A.J.;
RT "Germ-line mutations of the RET proto-oncogene in multiple endocrine
neoplasia type 2A.";
RL Nature 363:458-460(1993).
RN [8]
RP VARIANTS HSCR P-40; L-399; Q-762; P-765; Q-897; G-972 AND L-973.
RX MEDLINE=95219414; PubMed=7704557;
RA Yin L., Barone V., Sri M., Bolino A., Bocciardi R., Ceccherini I.,
RA Fasini B., Tocco I., Lerone M., Cywes S., Moore S.,
RA Vanderwinden J.-M., Abramowicz M.J., Kristoffersson U., Larsson L.T.,
RA Hamel B.C.J., Silengo M., Martucciello G., Romeo G.;
RT "Heterogeneity and low detection rate of RET mutations in Hirschsprung
disease.";
RL Eur. J. Hum. Genet. 2:272-280(1994).
RN [9]
RP VARIANT MEN2B THR-918.
RX MEDLINE=94272459; PubMed=7911697;
RA Eng C., Smith D.P., Mulligan L.M., Nagai M.A., Healey C.S.,
RA Ponder M.A., Gardner E., Scheumann G.F., Jackson C.E., Tunnacliffe A.,
RA Ponder B.A.J.;
RT "Point mutation within the tyrosine kinase domain of the RET
proto-oncogene in multiple endocrine neoplasia type 2B and related
sporadic tumours.";
RL Hum. Mol. Genet. 3:237-241(1994).
RN [10]

RP VARIANTS FMTC/MEN2A R-618; S-618; F-620; R-620; F-634; G-634 AND
Y-634.
RX MEDLINE=94348513; PubMed=7915165;
RA Xue F., Yu H., Maurer L.H., Menoli V.A., Nutile-Mcmenemy N.,
RA Schuster M.K., Browden D.W., Mao J.-I., Noll W.W.;
RT "Germline RET mutations in MEN 2A and FMTC and their detection by
simple DNA diagnostic tests.";
RL Hum. Mol. Genet. 3:635-638(1994).
RN [11]
RP VARIANTS FMTC/MEN2A TYR-609; ARG-618; SER-618 AND SER-620.
RX MEDLINE=95152521; PubMed=7849720;
RA Blaugrund J.E., Johns M.M. Jr., Eby Y.J., Ball D.W., Baylin S.B.,
RA Ruban R.H., Sidransky D.;
RT "RET proto-oncogene mutations in inherited and sporadic medullary
thyroid cancer.";
RL Hum. Mol. Genet. 3:1895-1897(1994).
RN [12]
RP VARIANTS FMTC, AND VARIANTS MEN2A.
RX MEDLINE=95179108; PubMed=7874109;
RA Schuffenecker I., Billaud M., Calender A., Chambe B., Ginot N.,
RA Calmettes C., Modigliani E., Lenoir G.M.;
RT "RET proto-oncogene mutations in French MEN 2A and FMTC families.";
RL Hum. Mol. Genet. 3:1939-1943(1994).
RN [13]
RP VARIANT HSCR W-609, VARIANT HSCR/MEN2A R-618, AND VARIANT HSCR/FMTC
R-620.
RX MEDLINE=95187155; PubMed=7881414;
RA Mulligan L.M., Eng C., Attie T., Lyonnet S., Marsh D.J., Hyland V.J.,
RA Robinson B.G., Frilling A., Verellen-Dumoulin C., Safar A.,
RA Venter D.J., Munnich A., Ponder B.A.J.;
RT "Diverse phenotypes associated with exon 10 mutations of the RET
proto-oncogene.";
RL Hum. Mol. Genet. 3:2163-2167(1994).
RN [14]
RP VARIANT MEN2B THR-918.
RX MEDLINE=94159102; PubMed=7906866;
RA Hofstra R.M.W., Landsvater R.M., Ceccherini I., Stulp R.P.,
RA Stelwagen T., Luo Y., Pasini B., Hoepfner J.W.M., van Amstel H.K.P.,
RA Romeo G., Lips C.J.M., Buys C.H.C.M.;
RT "A mutation in the RET proto-oncogene associated with multiple
endocrine neoplasia type 2B and sporadic medullary thyroid
carcinoma.";
RL Nature 367:375-376(1994).
RN [15]
RP VARIANTS HSCR PRO-765; GLN-897 AND GLY-972.
RX MEDLINE=94159103; PubMed=8114938;
RA Romeo G., Ronchetto P., Luo Y., Barone V., Seri M., Ceccherini I.,
RA Pasini B., Bocciardi R., Lerone M., Kaarainen H., Martucciello G.;
RT "Point mutations affecting the tyrosine kinase domain of the RET
proto-oncogene in Hirschsprung's disease.";
RL Nature 367:377-378(1994).
RN [16]
RP VARIANTS HSCR LEU-32; LEU-64; GLN-330 AND LEU-393.
RX MEDLINE=94159104; PubMed=8114939;
RA Edery P., Lyonnet S., Mulligan L.M., Pelet A., Dow E., Abel L.,
RA Holder S., Nihoul-Pkete C., Ponder B.A.J., Munnich A.;
RT "Mutations of the RET proto-oncogene in Hirschsprung's disease.";
RL Nature 367:378-380(1994).
RN [17]
RP VARIANT MEN2B THR-918.
RX MEDLINE=94151373; PubMed=7906417;
RA Carlson K.M., Dou S., Chi D., Scavarda N., Toshima K., Jackson C.E.,
RA Wells S.A. Jr., Goodfellow P.J., Donis-Keller H.;
RT "Single missense mutation in the tyrosine kinase catalytic domain of
the RET protooncogene is associated with multiple endocrine neoplasia
type 2B.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:1579-1583(1994).
RN [18]
RP VARIANTS MTC; FMTC; MEN2A AND MEN2B.
RX MEDLINE=96223053; PubMed=8625130;
RA Komminoth P., Kunz E.K., Matias-Guiu X., Hiort O., Christiansen G.,
RA Colomer A., Roth J., Heitz P.U.;
RT "Analysis of RET protooncogene point mutations distinguishes heritable


```
DR PRINTS; PR00452; SH3DOMAIN.
DR PRINTS; PR00402; TECBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Euk_pkinase; 1.
DR PRODOM; PD000066; SH3; 1.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TyrcK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00001; SH2; 1.
DR PROSITE; PS00002; SH3; 1.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR Transferase; Tyrosine-protein kinase; ATP-binding; SH2 domain;
KW SH3 domain; Phosphorylation; PH.
FT DOMAIN 4 111
FT DOMAIN 178 238
FT DOMAIN 246 344
FT DOMAIN 369 622
FT NP_BIND 375 383
FT BINDING 397 397
FT ACT_SITE 488 488
FT MOD_RES 518 518
FT VARSPPLIC 1 94
EGRARKKYKGVVDISKIKVEIVKNDGIVPCQKPFQV
VHDANTLYIFAP -> MNVSPFKINFS (IN ISOFORM
2).
MISSING (IN ISOFORM 2).
RPEGRPSLELLRTIDELVECEETFGF -> ESCLCRVAQD
LSSKNLIGSRF (IN ISOFORM 3).
V -> E (IN REF. 3).
FGVL -> YGIP (IN REF. 3).
L -> F (IN REF. 2).
SEQUENCE 630 AA; 73456 MW; 262640EE90D4A6D2 CRC64;
Query Match 24.1%; Score 301.5; DB 1; Length 630;
Best Local Similarity 32.4%; Pred. No. 5.6e-21;
Matches 71; Conservative 44; Mismatches 89; Indels 15; Gaps 7;
QY 2 ELRQLSRVNHPNVLKLYGACL--NPVCLVMEYAEAGSLVNLVHGAEPLPYVTAHAMSGC 59
Db 412 EAKVMKLTPLKVLQVLYGVTQOKPIYVTEMERGCLLNFILRQK--CHFSRDLMSWC 469
QY 60 LQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGGTVLKICDFGTA----CDIQTHMTNKK 115
Db 470 QDVCEGMEYL--ERNSTFHRDLAARNCLVNEAG-VVKVSDFGMARYVLDDQYTSSSGAK 525
QY 116 GSAAMWAPVEFGSNYSKCDVFSWGIILLWEVITR-RKPFDEIGGPAFRIMWVHNGTRP 174
Db 526 FPKVWCPEPEVFNYSFSSKSDVNSFGVLWMEIFTGGRMPFEK--NTNVEVVTMTVRGHLR 583
QY 175 PLIKLNPKPESLMTKCSWKDPQSRPSMEIVKIMPHLM 213
Db 584 HRPKLATKVLVEWLRQWERPEGRPSLEDLLRTIDELV 622
RESULT 14
TEC_HUMAN
ID TEC_HUMAN STANDARD; PRT; 631 AA.
AC P42680;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase Tec (EC 2.7.1.112).
GN TEC OR BSCRK4
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
```



```
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00779; BTK; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRINTS; PR00402; TECBTKDOMAIN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR ProDom; PD000086; SH3; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00107; BTK; 1.
DR SMART; SM00233; PH; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00219; TYKFC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00109; PROTEIN KINASE TYR; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50002; SH3; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH2 domain; SH3 domain; 3D-structure.
FT DOMAIN 4 117 PH.
FT DOMAIN 177 237 SH3.
FT DOMAIN 245 343 SH2.
FT DOMAIN 368 620 PROTEIN KINASE.
FT NP_BIND 374 382 ATP (BY SIMILARITY).
FT BINDING 396 396 ATP (BY SIMILARITY).
FT ACT_SITE 487 487 BY SIMILARITY.
FT MOD_RES 517 517 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 82 87 MISSING (IN REF. 2, 3 AND 4).
FT CONFLICT 535 535 F -> S (IN REF. 3).
FT CONFLICT 540 540 Y -> C (IN REF. 3).
SQ SEQUENCE 625 AA; 72291 MW; F7A4A18A8A1AADDG CRC64;

Query Match 23.9%; Score 299.5; DB 1; Length 625;
Best Local Similarity 32.1%; Pred. No. 8.6e-21;
Matches 72; Conservative 37; Mismatches 88; Indels 27; Gaps 9;

QY 2 ELROLSRVNHPNIVKLYGACLN--PVCVMVEYAEQGGSLYNVLHGAEPLPYVYTAHAMSWC 59
Db 411 EAEVMMKLSHPKLVQLYGVCLQEQAPICLVFEFWEHGCGLSDYLSQRGL--FAAETLLGMC 468
QY 60 LQCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDPG----TACDIQTHMTNNK 115
Db 469 LDVCEGMAYL--EKACVIRHDLAARN-CLVGENQVIKVSDFGWTFRVLDQYTSSTGTK 524
QY 116 GSAAWMAPEVPEGNSYSEKCDVFSWGILLWEVITRRK-PFDEIGGPAFRIMWAVHNGTRP 174
Db 525 FPKWASPEVFSFSGYSKSDVMSFGVLMWEVFESEKIPYENRSNS--EVVEDISTGFR- 581
QY 175 PLIKNLKP-----IESLMTRCWSKDPSPORPSMEEIVKIMTHL 212
Db 582 -----LYKPLASCHVYQIMNHCWKKEPEDRPPFSQLSLAEI 620
```

Search completed: December 10, 2002, 03:57:41
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:51:50 ; Search time 31 Seconds
(without alignments)

1515.443 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHPIVKLYGA.....MTHLMRYFPFGADEPLQLYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.rvirus.*
16: sp.bacteriaph.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	478	4 Q9NTR4	Q9ntr4 homo sapien
2	1252	100.0	491	4 Q9NZ70	Q9nz70 homo sapien
3	1252	100.0	539	4 Q9NTR1	Q9ntr1 homo sapien
4	1252	100.0	566	4 Q9NTR2	Q9ntr2 homo sapien
5	1252	100.0	606	11 Q923A8	Q923a8 mus musculus
6	1221	97.5	616	13 Q923A8	Q923a8 mus musculus
7	700	55.9	678	5 Q9V3Q6	Q9v3q6 drosophila
8	414	33.1	252	5 Q9VCV0	Q9vcv0 drosophila
9	376	30.0	371	13 Q9QZY8	Q9qzy8 brachydanio
10	374	29.9	411	10 Q9ZQ31	Q9zq31 arabidopsis
11	373	29.8	412	10 Q9M085	Q9m085 arabidopsis
12	372	29.7	454	11 Q9ESL3	Q9esl3 mus musculus
13	372	29.7	802	11 Q9ESL4	Q9esl4 mus musculus
14	371	29.6	455	4 Q9HCC4	Q9hcc4 homo sapien
15	371	29.6	800	4 Q9HDD2	Q9hdd2 homo sapien
16	371	29.6	800	4 Q9HCC5	Q9hcc5 homo sapien

17	371	29.6	800	4 Q9NYL2	Q9nyl2 homo sapien
18	371	29.6	800	4 Q9NYB9	Q9nyb9 homo sapien
19	368	29.4	416	10 Q94C42	Q94c42 triticum ae
20	366	29.2	637	10 Q94J41	Q94j41 oryza sativ
21	363.5	29.0	1020	5 Q9W313	Q9w313 drosophila
22	363.5	29.0	1148	5 Q9SVF6	Q9svf6 drosophila
23	363.5	29.0	1161	5 Q9SUN8	Q9sun8 drosophila
24	362	28.9	422	5 Q23846	Q23846 dictyosteli
25	362	28.9	1338	5 Q23927	Q23927 dictyosteli
26	360.5	28.8	1065	4 Q9H2N5	Q9h2n5 homo sapien
27	356	28.4	462	10 Q39886	Q39886 glycine max
28	355.5	28.4	370	10 Q9S7D5	Q9s7d5 arabidopsis
29	355	28.4	564	4 Q9H1Y7	Q9h1y7 homo sapien
30	353.5	28.2	859	4 Q8WY25	Q8wy25 homo sapien
31	351.5	28.1	977	5 Q9VW24	Q9vw24 drosophila
32	351	28.0	570	4 Q8WNN2	Q8wnn2 homo sapien
33	351	28.0	1036	4 Q8WNN1	Q8wnn1 homo sapien
34	350	28.0	546	10 Q22558	Q22558 arabidopsis
35	349	27.9	1001	11 Q8VDG6	Q8vdg6 mus musculu
36	348	27.8	391	10 Q22100	Q22100 arabidopsis
37	347.5	27.8	850	11 Q9DJ15	Q9dj15 mus musculu
38	347	27.7	345	10 Q9FGS7	Q9fgs7 arabidopsis
39	347	27.7	475	10 Q9STG5	Q9stg5 arabidopsis
40	346	27.6	570	10 Q8RWL6	Q8rw16 arabidopsis
41	346	27.6	740	5 Q21982	Q21982 caenorhabdi
42	346	27.6	1030	10 Q9C9U5	Q9c9u5 arabidopsis
43	345.5	27.6	553	10 Q81808	Q81808 arabidopsis
44	345.5	27.6	847	4 Q16584	Q16584 homo sapien
45	341.5	27.3	405	10 Q9LVQ9	Q9lvq9 arabidopsis

ALIGNMENTS

RESULT 1
Q9NTR4
ID Q9NTR4 PRELIMINARY; PRT; 478 AA.
AC Q9NTR4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE DJ154G14.1.3 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1c (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracev A;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87604.1; -.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS00011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
KW Kinase; Serine/threonine-protein kinase.
FT NON TER 1
SQ SEQUENCE 478 AA; 52482 MW; 177CC8CFA8D8DBF8 CRC64;
Query Match 100.0%; Score 1252; DB 4; Length 478;
Best Local Similarity 100.0%; Pred. No. 1.5e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 VELRQLSRVNHNPNIIVKLYGACLNPCVLWMEYAEAGSLIYNVLHGAEPLPYTTAAHAMSACL 60
DB 36 VELRQLSRVNHNPNIIVKLYGACLNPCVLWMEYAEAGSLIYNVLHGAEPLPYTTAAHAMSACL 95
QY 61 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGTACDIQTHMTNKKSAAM 120
DB 96 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGTACDIQTHMTNKKSAAM 155
QY 121 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 156 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 215
QY 181 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 216 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 263
RESULT 2
Q9NZ70 PRELIMINARY; PRT; 491 AA.
ID Q9NZ70; Q9NZ70;
AC Q9NZ70; Q9NZ70;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE TGF-beta-activated kinase splice variant d (DJ154G14.1.4)
DE (mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1)))
DE GN TAK1 OR MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20568765; PubMed=11118615;
RA Dempsey C.E., Sakurai H., Sugita T., Guesdon F.;
RT "Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1."
RT Biochim. Biophys. Acta 1517:46-52 (2000).
RN [2]
RP SEQUENCE OF 41-491 FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AF218074; AAF27652.1; -.
DR EMBL; AL121964; CAB87605.1; -.
DR HSSP; P08631; 1ADM.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Kinase; Serine/threonine-protein kinase; Transferrase.
KW SEQUENCE 491 AA; 53739 MW; B7D8832E286A99C5 CRC64;
SQ
Query Match 100.0%; Score 1252; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.6e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
DB 136 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGTACDIQTHMTNKKSAAM 195
QY 121 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 196 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 255
QY 181 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 256 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 303
RESULT 3
Q9NTR1 PRELIMINARY; PRT; 539 AA.
ID Q9NTR1; Q9NTR1;
AC Q9NTR1; Q9NTR1;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE DJ154G14.1.1 (mitogen-activated protein kinase kinase kinase 7 (TGF-beta activated kinase 1a (TAK1))) (Fragment).
DE GN MAP3K7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AL121964; CAB87607.1; -.
DR HSSP; P12931; 1PMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYK; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
DR PROSITE; PS0011; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR Kinase; Serine/threonine-protein kinase.
KW NON-TER 1
RP SEQUENCE 539 AA; 60006 MW; B6183F53CC7F324 CRC64;
SQ
Query Match 100.0%; Score 1252; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 1.7e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VELRQLSRVNHNPNIIVKLYGACLNPCVLWMEYAEAGSLIYNVLHGAEPLPYTTAAHAMSACL 60
DB 36 VELRQLSRVNHNPNIIVKLYGACLNPCVLWMEYAEAGSLIYNVLHGAEPLPYTTAAHAMSACL 95
QY 61 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGTACDIQTHMTNKKSAAM 120
DB 96 QCSGVAAYLHSMQPKALIHRLDKPNNLLVAGTGLKICDFGTACDIQTHMTNKKSAAM 155
QY 121 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 180
DB 156 MAPEVEGSSNSSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMAVHNGTRPPLIKNL 215
QY 181 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 228
DB 216 PKPIESLMTRCMSKDPSPSRPMEIIVKIMTHLMRYFPGADEPLQYPCQ 263
RESULT 4
Q9NTR2 PRELIMINARY; PRT; 566 AA.
ID Q9NTR2; Q9NTR2;
AC Q9NTR2; Q9NTR2;
DT 01-OCT-2000 (Tremblrel. 15, Created)
```

```

DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE DJ154G14.1.2 (Mitogen-activated protein kinase kinase 7
DE (TGF-beta activated kinase 1b (TAK1))) (Fragment).
GN MAP3K7.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Tracey A.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBSJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC EMBL; AL121964; CAB87606.1; -.
DR HSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase; Serine/threonine-protein_kinase.
FT NON_TER 1
SQ SEQUENCE 566 AA; 62972 MW; 32CDAC1211B200CF CRC64;

Query Match 100.0%; Score 1252; DB 4; Length 566;
Best Local Similarity 100.0%; Pred. No. 1.8e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 60
Db 36 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 95
Qy 61 QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
Db 96 QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 155
Qy 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 156 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 215
Qy 181 PKPIESLMTRCWSKDPQRSMSBEIVKIMTHLMRYFFGADEPLQYPCQ 228
Db 216 PKPIESLMTRCWSKDPQRSMSBEIVKIMTHLMRYFFGADEPLQYPCQ 263

RESULT 5
Q923A8 ID Q923A8 PRELIMINARY; PRT; 606 AA.
AC Q923A8;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5989).
DR MAP3K7.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC006665; AAH06665.1; -.
DR MGD; MGI:1346877; Map3k7.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.

```

```

DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; UNKNOWN 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN 1.
KW ATP-binding; Transferase.
SQ SEQUENCE 606 AA; 67194 MW; AB8664F389272102 CRC64;

Query Match 100.0%; Score 1252; DB 11; Length 606;
Best Local Similarity 100.0%; Pred. No. 2e-118;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 60
Db 76 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 135
Qy 61 QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 120
Db 136 QCSQGVAYLHSMQPKALIHRLKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW 195
Qy 121 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 180
Db 196 MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL 255
Qy 181 PKPIESLMTRCWSKDPQRSMSBEIVKIMTHLMRYFFGADEPLQYPCQ 228
Db 256 PKPIESLMTRCWSKDPQRSMSBEIVKIMTHLMRYFFGADEPLQYPCQ 303

RESULT 6
O73613 ID O73613 PRELIMINARY; PRT; 616 AA.
AC O73613;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE TAK1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98130593; PubMed=9463380;
RA Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,
RA Matsumoto K., Nishida E., Ueno N.;
RT "Role of TAK1 and TAB1 in BMP signaling in early Xenopus
RT development.";
RL EMBL J. 17:1019-1028(1998).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; U92030; AAC14008.1; -.
DR HSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR004040; STY_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Euk_pkinase; 1.
DR SMART; SM00221; STYKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Serine/threonine-protein_kinase; Transferase.
SQ SEQUENCE 616 AA; 68464 MW; 493AD2A05ADC38B6 CRC64;

Query Match 97.5%; Score 1221; DB 13; Length 616;
Best Local Similarity 97.4%; Pred. No. 2.8e-115;
Matches 222; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 60
Db 65 VELQLSRVNHNPINVKLYGACLPVCLVMEYAEAGGSLYNVLHGAEPPLPYTTAAHAMSACL 124

```

```
QY 61 QCSQGVAYLHSMQPKALIHRLDKPRLLLVAGTGLKICDFGTACDIOTHTMTNKGSAW 120
DB 125 QCAQGAVALHSMKPKALIHRLDKPRLLLVAGTGLKICDFGTACDIOTHTMTNKGSAW 184
QY 121 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGPAFRIMAAVANGTRPPLIKL 180
DB 185 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGPAFRIMAAVANGTRPPLIKL 244
QY 181 PKPIESLMTRCWSKDSQRPSPMEIYKIMTHLMRYPPGADDEPLYQCC 228
DB 245 PKPIESLMTRCWSKDSQRPSPMEIYKIMTHLMRYPPGADVLSIQYPCQ 292

RESULT 7
Q9V306 PRELIMINARY; PRT; 678 AA.
ID Q9V306
AC Q9V306;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Putative TAK1 protein (CG1388 protein) (LD42274P).
GN TAK1 OR CG1388 OR CG18492.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RX STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandal D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo E., Delcher A.L., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel K., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulj D., Lai Z.,
RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skuski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang X., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Takatsu Y., Nakamura M., Stapleton M., Danos M., Matsumoto M.,
RA O'Connor M.B., Shibuya H., Ueno N.;
```

```
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Garin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacle J., Parag V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.B., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AEO03571; AAF050895.1; -
DR EMBL: AP199466; AAF06815.1; -
DR EMBL: AY051953; AAK93377.1; -
DR HSSP: P08631; 1AD5.
DR FlyBase: FBgn0026323; Tak1.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001990; Granlin.
DR InterPro: IPR002290; Ser_thr_pkinase.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00221; STYKC; 1.
DR PROSITE: PS00422; GRANIN_1; UNKNOWN_1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding; Hypothetical protein; Kinase;
KW Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 678 AA; 75674 MW; 87EBA80CDBBCD45 CRC64;

Query Match 55.3%; Score 700; DB 5; Length 678;
Best Local Similarity 57.5%; Pred. No. 2..2e-62;
Matches 130; Conservative 35; Mismatches 57; Indels 4; Gaps 3;

QY 2 ELRQLSRNHNHYLYKG--ACLNPCVLNMEYAGGSLYVHLGAEPLPYTAHMSWC 59
DB 60 EVKLLSRKHPNIIIALGISSYQCATYLIWFAGGSLHNFHLG-KVPAYSLAHMSWA 118
QY 60 LQCSQGVAYLHSMQPKALIHRLDKPRLLLVAGTGLKICDFGTACDIOTHTMTNKGSAW 119
DB 119 RQCBGLAYLHAMPKPLIHRLDVKPLNLLTNKGRNLKICDFGTADKSTMTNKGSAW 178
QY 120 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGPAFRIMAAVANGTRPPLIKN 179
DB 179 MAEEVEGNSYSEKCDVFSWGIIILMEVITRRKPFDEIGGPAFRIMAAVANGTRPPLITT 237
QY 180 PKPIESLMTRCWSKDSQRPSPMEIYKIMTHLMRYPPGADDEPLYQ 225
DB 228 PKPIEDLMTACWKTVPEDRPSMOYIVGVNHEIYKDYTGADKALEY 283

RESULT 8
Q9VCV0 PRELIMINARY; PRT; 252 AA.
ID Q9VCV0
AC Q9VCV0;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE CG4803 protein.
GN CG4803.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxID=7227;
RX NCB1_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA ADAMS M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
```



```

RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Gutter R.A., Lewis S.E., Richards S., Asburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolischmann W.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Chery J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodzik A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Haridas N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinart K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong X., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "the genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195 (2000).
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL: AF003740; AAF56055.1; -.
DR FlyBase: FBgn00374015; CG4803.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR002230; Ser_thr_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR ProDom: PD000001; Euk_pkinase; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; 1.
DR ATP-binding: Serine/threonine-protein kinase, Transferase.
DR KW SEQUENCE 252 AA; 28853 MW; 8E006F2EAB68D17 CRC64;
SQ
Query Match 33.1%; Score 414; DB 5; Length 252;
Best Local Similarity 41.8%; Pred. No. 6.4e-34;
Matches 82; Conservative 35; Mismatches 53; Indels 26; Gaps
QY 2 ELRQLSRVNHPTIKVLYGACLNIVC--LYMVEAEGSLYVNLHGAEPLUPYTAHAMSWC 59
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 54 EIVQLTKASHWNIVELYGTSRHEGCGALLMEFVDGGLSSFLH-AKSPKSYSHAHFNWA 112
QY 60 LQCSQGVAYLHSHQPKALIHRLDKPNLLVAGGTGVLKICDFGTACDQTHMTNKGSA 119
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 113 HOIAQGIAYLHGQPKAVIHRLDKPLNTLLCEKGLKLIKICDFGTVDVLSQSISNAGTCR 172
QY 120 WMAPE-----VPEGSNYSKCDQVFSNGIILNEVITRRKPFDEI 157
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 173 YKAPEVRELFFDKSNRIINQPTGFKVQLQGNPKDEKCDVSNWITFWELLSRKEPFQY 232
QY 158 GGPAPRIMWAVHNGTR 173
| : : : : : | : : : : : |
Db 233 -NTLFELYMINEGKR 247
RESULT 9
Q90ZY8

```

RA Tang C., Toriumi M., Yu G., Yu S., Bowser L., Carinici P., Chen H.,
 RA Cheuk R., Hayashizaki Y., Ishida J., Jones T., Kamya A., Lin J.,
 RA Karlin-Neumann G., Kawai J., Kim C., Koesema E., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shin P., Southwick A., Tracy S.E.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT "Full length cDNA of gene T8124.9/At2g24360 (GI:4337195).";
 RT Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RL EMBL; AC0064403; M18109.2; -;
 DR EMBL; AY046026; AAK76700.1; -;
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; UNKNOWN; 1.
 DR Hypochemical protein; ATP-binding; Transferase.
 KW SEQUENCE 411 AA; 46001 MW; 387001CFBB411BAA CRC64;
 SQ

Query Match 29.9%; Score 374; DB 10; Length 411;
 Best Local Similarity 39.4%; Pred. No. 1.3e-29;
 Matches 86; Conservative 40; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNL---HGAEPLPYTAHA 55
 DB 179 EVSMLANLKHPIVIRFGACIKPVMWCIVTEYAKGGSVRFTRQRNAVPLKL-----A 233
 QY 56 MSWCLQSGVAYVHSMQPKALIHRLDKPNNLLVAGGTVLKI CDFTG-CDIQTH-MTN 113
 DB 234 VKQALDVARGAAYVHG---RNFHRLDKSDNLLISADKSI-KIADFGVARIIEVQTEGMTP 289
 QY 114 NKSAAMAAPEVEFGSNYSKCVFSWGIIIMVEVITRRKPFDEIGG--PAFRIMAVVHG 171
 DB 290 ETGYRMAAPEMIQHRPYTKVDVYFGIVLWELITGLPFOQMTAVQAFAV---VNRG 346
 QY 172 TRPPLIKNLPKPIESLMTRCMSKDPSPQPSMEIYKIM 209
 DB 347 VRPTVPADCLPVLGEMITRCMDADPEVRPCFAEIVNL 385

RESULT 11
 Q9M085 PRELIMINARY; PRT; 412 AA.
 AC Q9M085;
 DT 01-OCT-2000 (Tremblrel. 15, Created)
 DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE Protein kinase-like protein.
 GN AT4G31170
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroside II; Brassicales; Brassicaceae; Arabidopsi.
 OX NCBI_Taxid=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Lennard N., Quail M., Harris B., Rajandream M.A., Barrell B.G.,
 RA Mewes H.W., Lemcke K., Mayer K.F.X.,
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AL161578; CAB79835.1; -;
 DR HSSP; P08631; IAD5.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; TYR_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; SM00219; TYTKC; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR ProDom; PD000001; Euk_pkinase; 1.

DR SMART; SM00221; STYKc; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 412 AA; 46083 MW; 644F35A90210D48 CRC64;
 Query Match 29.8%; Score 373; DB 10; Length 412;
 Best Local Similarity 39.9%; Pred. No. 1.7e-29;
 Matches 87; Conservative 39; Mismatches 70; Indels 22; Gaps 9;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNL---HGAEPLPYTAHA 55
 DB 180 EVSMLANLKHPIVIRFGACIKPVMWCIVTEYAKGGSVRFTRQRNAVPLKL-----A 234
 QY 56 MSWCLQSGVAYVHSMQPKALIHRLDKPNNLLVAGGTVLKI CDFTG-CDIQTH-MTN 113
 DB 235 VKQALDVARGAAYVHG---RNFHRLDKSDNLLISADKSI-KIADFGVARIIEVQTEGMTP 290
 QY 114 NKSAAMAAPEVEFGSNYSKCVFSWGIIIMVEVITRRKPFDEIGG--PAFRIMAVVHG 171
 DB 291 ETGYRMAAPEMIQHRPYTKVDVYFGIVLWELITGLPFOQMTAVQAFAV---VNRG 347
 QY 172 TRPPLIKNLPKPIESLMTRCMSKDPSPQPSMEIYKIM 209
 DB 348 VRPTVPADCLPVLGEMITRCMDADPEVRPCFAEIVNL 385

RESULT 12
 Q9ESL3 PRELIMINARY; PRT; 454 AA.
 AC Q9ESL3;
 DT 01-MAR-2001 (Tremblrel. 16, Created)
 DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
 DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
 DE MLTK-beta.
 GN ZAK OR MLTK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=2164927; Pubmed=11042189;
 RA Gotch I., Adachi M., Nishida E.,
 RT "Identification and Characterization of a Novel MAP Kinase Kinase
 RT Kinase, MLTK.";
 RL J. Biol. Chem. 276:4276-4286(2001).
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; AB049732; BAB16443.1; -;
 DR HSSP; P12931; IPMK.
 DR MGD; MGI:1931274; Zak.
 DR InterPro; IPR000719; Euk_pkinase.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR004040; STY_pkinase.
 DR InterPro; IPR001245; TYR_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Euk_pkinase; 1.
 DR SMART; SM00221; STYKc; 1.
 DR SMART; SM00220; S_TKc; 1.
 DR SMART; SM00219; TYTKC; 1.
 DR PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 454 AA; 51366 MW; 35C2FCD0729D9395 CRC64;
 Query Match 29.7%; Score 372; DB 11; Length 454;
 Best Local Similarity 37.7%; Pred. No. 2.4e-29;
 Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLPV--CLVMEYAGGSLYVNL---HGAEPLPYTAHA 57
 DB 53 EALILSVLSHKNITQFYGVILEPENVIGVTEYASLSGLYINSNSEE---DMEHIMT 109

Search completed: December 10, 2002, 03:58:25
Job time : 33 secs

```
QY 58 WCLQCSQGVAYIHSMPKALIHRLKPPNLLVAGSTVIKIDFGTACDIQTHMTNKK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVYIHRDLKSRNVVIAADG-VLKIDFG-ASRFNHTTHMSLV 167
QY 116 GSAAMWAPPEVFGSNTSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIMW-AVANGTRP 174
DB 168 GTFPMWAPPEVIOQLPVSETCDTYSYGVVLMEMLTREVPPKGLGG--LQVAMLVKERNRL 225
QY 175 PLIKNLPRPESIMTRCWSKDPSPQSPMEIVKIM 209
DB 226 TTPSSCRPSFAELLHQCWEADAKKRPSPFOIISIL 260

RESULT 15
Q9HDD2 PRELIMINARY; FRT; 800 AA.
ID Q9HDD2
AC Q9HDD2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Plaucible mixed-lineage kinase protein.
GN MLKLAK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphoid Organ;
RA Abe Y, Ueda N;
RT "Plaucible mixed-lineage kinase derived from LAK cell.";
RL Submitted (JUL-1999) to the EMBL/Genbank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR EMBL; AB030034; BAB12040.1; -.
DR HSSP; P12931; 1FMK.
DR InterPro; IPR000719; Euk_kinase.
DR InterPro; IPR001660; SAM.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR004040; STY_kinase.
DR InterPro; IPR01245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00536; SAM; 1.
DR PRINTS; PRO0109; TYRKINASE.
DR ProDom; PD000001; Euk_kinase; 1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00221; STYK; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
SQ SEQUENCE 800 AA; 9115 MW; B2814509EC54B07A CRC64;

Query Match 29.6%; Score 371; DB 4; Length 800;
Best Local Similarity 37.7%; Pred.No.6e-29;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACINP--VCLVMEVYEGSLYVNLHG--AEPLPYTTAAHMS 57
DB 53 EAILSLVLSHRNIIOFYGVLEBPNGIYTEVYSLGSLVDYINSNRSEEM--DMDHIMT 109
QY 58 WCLQCSQGVAYIHSMPKALIHRLKPPNLLVAGSTVIKIDFGTACDIQTHMTNKK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVYIHRDLKSRNVVIAADG-VLKIDFG-ASRFNHTTHMSLV 167
QY 116 GSAAMWAPPEVFGSNTSEKCDVFSWGIILMEVITRRKPPDEIGGPAFRIMW-AVANGTRP 174
DB 168 GTFPMWAPPEVIOQLPVSETCDTYSYGVVLMEMLTREVPPKGLGG--LQVAMLVKERNRL 225
QY 175 PLIKNLPRPESIMTRCWSKDPSPQSPMEIVKIM 209
DB 226 TTPSSCRPSFAELLHQCWEADAKKRPSPFOIISIL 260
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:56:30 ; Search time 19 Seconds
(without alignments)
353.075 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELQLSRVNHNPVILVLYGA.....MTHLMRYPPGADEPLQVPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/prodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/prodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/prodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/prodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/prodata/1/iaa/backfiles.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	579	US-09-529-279-4	Sequence 4, Appli
2	1252	100.0	590	US-09-529-279-15	Sequence 15, Appli
3	371	29.6	455	US-09-221-235-5	Sequence 5, Appli
4	371	29.6	455	US-09-221-928-5	Sequence 5, Appli
5	371	29.6	455	US-09-221-527-5	Sequence 5, Appli
6	371	29.6	455	US-09-221-236-5	Sequence 5, Appli
7	371	29.6	455	US-09-221-416-5	Sequence 5, Appli
8	371	29.6	455	US-09-221-245-5	Sequence 5, Appli
9	371	29.6	455	US-09-163-115-5	Sequence 5, Appli
10	371	29.6	455	US-09-221-528-5	Sequence 5, Appli
11	371	29.6	455	US-09-593-553-5	Sequence 5, Appli
12	371	29.6	455	US-09-221-237-5	Sequence 5, Appli
13	353.5	28.2	668	US-08-205-018-2	Sequence 2, Appli
14	353.5	28.2	859	US-08-395-580-2	Sequence 2, Appli
15	353.5	28.2	859	PCT-US95-02792-2	Sequence 2, Appli
16	338.5	27.0	835	US-09-291-839-2	Sequence 2, Appli
17	323	25.8	821	US-07-928-464-2	Sequence 2, Appli
18	323	25.8	821	US-08-003-311B-2	Sequence 2, Appli
19	323	25.8	821	US-08-261-432-2	Sequence 2, Appli
20	323	25.8	821	PCT-US93-07347-2	Sequence 2, Appli
21	320	25.6	263	US-09-035-706-5	Sequence 5, Appli
22	320	25.6	263	US-08-955-841-5	Sequence 5, Appli
23	320	25.6	263	US-09-390-425-5	Sequence 5, Appli
24	320	25.6	263	US-09-566-906-5	Sequence 5, Appli
25	309.5	24.7	269	US-07-857-224B-79	Sequence 79, Appli
26	307.5	24.6	276	US-07-857-224B-72	Sequence 72, Appli
27	306	24.4	275	US-07-857-224B-71	Sequence 71, Appli

28	305.5	24.4	787	4	US-09-188-930-334	Sequence 334, App
29	304.5	24.3	304	2	US-08-701-191A-27	Sequence 27, Appl
30	302.5	24.2	316	1	US-08-278-089A-16	Sequence 16, Appl
31	302.5	24.2	316	2	US-08-838-957A-15	Sequence 15, Appl
32	299.5	23.9	527	4	US-08-426-509A-10	Sequence 10, Appl
33	299.5	23.9	527	5	PCT-US95-05008-10	Sequence 10, Appl
34	299.5	23.9	625	1	US-08-391-615-3	Sequence 3, Appli
35	298.5	23.8	620	4	US-08-426-509A-9	Sequence 9, Appli
36	298.5	23.8	620	5	PCT-US95-05008-9	Sequence 26, Appl
37	294.5	23.5	293	2	US-08-701-191A-26	Sequence 185, App
38	294	23.5	536	4	US-09-188-930-185	Sequence 55, Appl
39	293.5	23.4	274	2	US-08-469-537A-55	Sequence 27, Appl
40	290	23.2	1584	4	US-09-457-040B-27	Sequence 16, Appl
41	284.5	22.7	293	2	US-08-701-191A-16	Sequence 4, Appli
42	280	22.4	388	1	US-08-070-165F-4	Sequence 5, Appli
43	280	22.4	388	2	US-08-885-418-4	Sequence 5, Appli
44	278.5	22.2	588	1	US-08-391-615-5	Sequence 13, Appl
45	278	22.2	299	2	US-08-701-191A-13	

ALIGNMENTS

RESULT 1

US-09-529-279-4
; Sequence 4, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-4

Query Match	100.0%	Score 1252;	DB 4;	Length 579;
Best Local Similarity	100.0%;	Pred. No. 3.2e-133;		
Matches 228;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	VELQLSRVNHNPVILVLYGACLNPNVCLVMEYAEAGGSLYNVLHGAEPLPYVYTAAHAMSWCL	60	
Db	76	VELQLSRVNHNPVILVLYGACLNPNVCLVMEYAEAGGSLYNVLHGAEPLPYVYTAAHAMSWCL	135	
Qy	61	QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	120	
Db	136	QCSQGVAYLHSMQPKALIHRLDKPPNLLVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	195	
Qy	121	MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL	180	
Db	196	MAPEVFGSNYSKCDVFSWGIIILWEVITRRKPPDEIGGPAFRIMWAVHNGTRPPLIKNL	255	
Qy	181	PKPIESLMTRCWSKDPSPSMEEIVKIMTHLMRYFFGADEPLQVPCQ	228	
Db	256	PKPIESLMTRCWSKDPSPSMEEIVKIMTHLMRYFFGADEPLQVPCQ	303	

RESULT 2

US-09-529-279-15
; Sequence 15, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:

APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 15
LENGTH: 590
TYPE: PRF
ORGANISM: Homo sapiens
US-09-529-279-15

Query Match 100.0%; Score 1252; DB 4; Length 590;
Best Local Similarity 100.0%; Pred. No. 3,3e-133;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 60
DB 76 VELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 135
QY 61 QCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKSAAM 120
DB 136 QCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNKSAAM 195
QY 121 MAEVEGSGVSEKCVFSGIILWEVITRRKPFDEIGGPARIMAVHNGTRPPLIKL 180
DB 196 MAEVEGSGVSEKCVFSGIILWEVITRRKPFDEIGGPARIMAVHNGTRPPLIKL 255
QY 181 PKRISLMTRCWSKDPSPRSMEEIVKIMTHLMRYFPGADPEIQYPCQ 228
DB 256 PKRISLMTRCWSKDPSPRSMEEIVKIMTHLMRYFPGADPEIQYPCQ 303

RESULT 3
US-09-221-235-5
Sequence 5, Application US/09221235
Patent No. 6043040
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,235
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRF
ORGANISM: Homo sapiens
US-09-221-235-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 57
DB 53 ELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 109
QY 58 WCLQCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADG-VLKICDFG-ASRFNHTTHMSLV 167
QY 116 GSAAMAPEVEGSGVSEKCVFSGIILWEVITRRKPFDEIGGPARIMAVHNGTRP 174

DB 168 GTFPMAPEVIGSLPVSFETCDTYSYGVLMEMLTREVPFGLEG--LQVAMLYVEKNERL 225
QY 175 PLINLKPPIESLMTRCWSKDPSPRSMEEIVKIM 209
DB 226 TIPSCPSFAELHQCWEADAKKRPFSKQIISIL 260

RESULT 4
US-09-221-928-5
Sequence 5, Application US/09221928
Patent No. 6121030
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,928
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRF
ORGANISM: Homo sapiens
US-09-221-928-5

Query Match 29.6%; Score 371; DB 3; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

QY 2 ELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 57
DB 53 ELRLSRVNHPIVLYGACLNPCVCLVMEYAEAGSLVNVHLGAPLPYTTAAHMSWCL 109
QY 58 WCLQCSQGVAVYHSMQKALIHRLDKPNNLLVAGGTVLKICDFGTACDIQTHMTNNK-- 115
DB 110 WATDVAKGMHYLHMEAPVKVIRHDLKSRNVVIAADG-VLKICDFG-ASRFNHTTHMSLV 167
QY 116 GSAAMAPEVEGSGVSEKCVFSGIILWEVITRRKPFDEIGGPARIMAVHNGTRP 174
DB 168 GTFPMAPEVIGSLPVSFETCDTYSYGVLMEMLTREVPFGLEG--LQVAMLYVEKNERL 225
QY 175 PLINLKPPIESLMTRCWSKDPSPRSMEEIVKIM 209
DB 226 TIPSCPSFAELHQCWEADAKKRPFSKQIISIL 260

RESULT 5
US-09-221-527-5
Sequence 5, Application US/09221527
Patent No. 6146832
GENERAL INFORMATION:
APPLICANT: Acton, Susan
TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
FILE REFERENCE: NMI-050
CURRENT APPLICATION NUMBER: US/09/221,527
CURRENT FILING DATE: 1998-12-28
EARLIER APPLICATION NUMBER: 09/163,115
EARLIER FILING DATE:
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 455
TYPE: PRF
ORGANISM: Homo sapiens
US-09-221-527-5

Query Match 29.6%; Score 371; DB 4; Length 455;
Best Local Similarity 37.7%; Pred. No. 1,4e-33;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;


```

; TITLE OF INVENTION: Sequences Encoding the Same and Methods Related Thereto
;
; NUMBER OF SEQUENCES: 4
;
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and Norris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch disk, 720 KB
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02792
; FILING DATE: herewith
; CLASSIFICATION:
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/205,018
; FILING DATE: 01-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Rebecca L. Ralph (formerly Gaumond)
; REGISTRATION NUMBER: 35,152
; REFERENCE/DOCKET NUMBER: CH-0488
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 859 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-02792-2

```

Query Match 28.2%; Score 353.5; DB 5; Length 859;
Best Local Similarity 34.7%; Pred. No. 3.2e-31;
Matches 75; Conservative 43; Mismatches 81; Indels 17; Gaps 8;

```

QY 2 ELRLSRVNHPIVVKLYGACLNLP--VCLVMEYAEAGSLYVNLHGAPLPYTTAAHAMSWC 59
DB 161 DIKHLKRLKHPNITTFKGYCTGAPCYCIMEFCAQGLYEVLRAGRPF--TPSLVDWS 217
QY 60 LQCSQGVAVLHSMQPKALIHRLDKPPNLLVAGTVLKI CDFTACDIQTHMTNKK--GS 117
DB 218 MGJAGCMNYLHLHK--IIHRDLKSPN-MLITYDVVKISDFGTSKELSDKSTKWSFAGT 273
QY 118 AAMMAPEVFEFGSNYSKCDVFSWGIILWEVITRRKRPDEIGGPAPRIMAV--HNGTRPPL 176
DB 274 VAMMAPEVIRNEPVSEKVDIWSFGVVLWMLTGEIPIKVDSSA--IIMGVGSNSLHLPV 331
QY 177 IKNLKPKIESLMTKWSKDPSCQRPMSMEIWKIMTHL 212
DB 332 PSSCPDGFKILLRQCMNSKRPNRPSFRQ--ILLHL 364

```

Search completed: December 10, 2002, 03:59:30
Job time : 21 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 10, 2002, 03:57:20 ; Search time 10 Seconds

(without alignments)
370.325 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHNPVIVKLYGA.....MTHLMRYFFCADEPLQYPCQ 228

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 103943 seqs, 16242309 residues

Total number of hits satisfying chosen parameters: 103943

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1252	100.0	579	9	US-10-158-895-4
2	1252	100.0	590	9	US-10-158-895-15
3	371	29.6	455	10	US-09-757-982-5
4	363.5	29.0	394	10	US-09-862-027-19
5	355	28.4	1036	12	US-10-014-882-2
6	346	27.6	328	10	US-09-862-027-18
7	338.5	27.0	835	10	US-09-947-199-2
8	335.5	26.8	835	10	US-09-947-199-8
9	331.5	26.5	966	10	US-09-771-161A-197
10	320	25.6	263	10	US-09-840-704-5
11	313.5	25.0	850	10	US-09-904-389-2
12	312	24.9	265	10	US-09-797-039-10
13	312	24.9	277	10	US-09-815-915-13
14	312	24.9	277	10	US-09-882-166-4
15	312	24.9	278	9	US-09-842-582-4
16	312	24.9	278	10	US-09-797-039-13
17	312	24.9	278	10	US-09-922-138-18
18	312	24.9	278	10	US-09-922-138-27
19	312	24.9	278	10	US-09-910-150-17

20	312	24.9	278	10	US-09-910-150-31	Sequence 31, Appl
21	308	24.6	278	10	US-09-815-915-16	Sequence 16, Appl
22	307	24.5	237	10	US-09-916-790-17	Sequence 17, Appl
23	304	24.3	273	10	US-09-922-138-11	Sequence 11, Appl
24	303	24.2	247	10	US-09-916-790-7	Sequence 7, Appl
25	303	24.2	271	10	US-09-799-875-28	Sequence 28, Appl
26	303	24.2	272	10	US-09-780-949-4	Sequence 4, Appl
27	303	24.2	272	10	US-09-910-150-27	Sequence 27, Appl
28	301.5	24.1	279	10	US-09-799-875-23	Sequence 23, Appl
29	299.5	23.9	527	10	US-09-977-269-10	Sequence 10, Appl
30	298.5	23.8	620	10	US-09-977-269-9	Sequence 9, Appl
31	296	23.6	272	10	US-09-815-915-14	Sequence 14, Appl
32	292.5	23.4	280	10	US-09-515-806-15	Sequence 15, Appl
33	289	23.1	425	10	US-09-828-313-29	Sequence 29, Appl
34	288.5	23.0	628	10	US-09-862-027-48	Sequence 48, Appl
35	282.5	22.6	627	10	US-09-862-027-47	Sequence 47, Appl
36	278	22.2	310	9	US-09-939-833-7	Sequence 7, Appl
37	278	22.2	310	10	US-09-939-754-7	Sequence 7, Appl
38	278	22.2	310	10	US-09-939-832-7	Sequence 7, Appl
39	278	22.2	764	10	US-09-925-302-714	Sequence 714, App
40	273.5	21.8	214	10	US-09-515-806-10	Sequence 10, Appl
41	271.5	21.7	239	10	US-09-797-039-12	Sequence 12, Appl
42	271.5	21.7	245	10	US-09-815-915-15	Sequence 15, Appl
43	270.5	21.6	212	10	US-09-834-496A-4	Sequence 4, Appl
44	270.5	21.6	214	10	US-09-860-352A-6	Sequence 6, Appl
45	269.5	21.5	507	10	US-09-977-269-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-4
; Sequence 4, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: TSUCHIYA, MASAYUKI
; APPLICANT: OHTOMO, TOSHIHIKO
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 579
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-4

Query Match 100.0%; Score 1252; DB 9; Length 579;
Best Local Similarity 100.0%; Pred. No. 2.7e-104;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPLVPYTTAAHMSWCL	60
Db	76	VELRQLSRVNHNPVIVKLYGACLNVPCLVMEYAEAGGSLYNVLHGAEPLVPYTTAAHMSWCL	135
Qy	61	QCSQGVAYLHMQPKALIHRLDKPPNLLIVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	120
Db	136	QCSQGVAYLHMQPKALIHRLDKPPNLLIVAGGTVLKICDFGTACDIQTHMTNNKGSAAW	195
Qy	121	MAPEVFEAGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	180
Db	196	MAPEVFEAGNSYSEKCDVFSWGIILMEVITRRKPFDEIGGPAFRIMWAVHNGTRPPLIKNL	255

Qy 181 PKPIESLMTRCWSKDPSPQSRPSMEIYKIMTHLMRYFPGADEPLQYPCQ 228
|||
Db 256 PKPIESLMTRCWSKDPSPQSRPSMEIYKIMTHLMRYFPGADEPLQYPCQ 303

RESULT 2
US-10-158-895-15

/ Sequence 15, Application US/10158895
/ Patent No. US20020158624A1
/ GENERAL INFORMATION:
/ APPLICANT: ONO, KOICHIRO
/ APPLICANT: OHTOMO, TOSHIHIKO
/ APPLICANT: TSUCHIYA, MASAYUKI
/ TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
/ FILE REFERENCE: 053466/0278
/ CURRENT APPLICATION NUMBER: US/10/158,895
/ PRIOR FILING DATE: 2002-06-03
/ PRIOR APPLICATION NUMBER: US/09/529,279
/ PRIOR FILING DATE: 2000-04-11
/ PRIOR APPLICATION NUMBER: PCT/JP98/04796
/ PRIOR FILING DATE: 1998-10-22
/ PRIOR APPLICATION NUMBER: JP 9/290188
/ PRIOR FILING DATE: 1997-10-22
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 15
/ LENGTH: 590
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-158-895-15

Query Match 100.0%; Score 1252; DB 9; Length 590;
Best Local Similarity 100.0%; Pred. No. 2,8e-104;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VELRLSRVNHPIVLYGACINPVCVMEYVAGSGSLVNVLHGAEPPLPYTTAAHMSWCL 60
|||
Db 76 VELRLSRVNHPIVLYGACINPVCVMEYVAGSGSLVNVLHGAEPPLPYTTAAHMSWCL 135
|||
Qy 61 QCSQGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDICTHMTNNKSAAM 120
|||
Db 136 QCSQGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDICTHMTNNKSAAM 195
|||
Qy 121 MADEVEGNSYSEKCVFSWGIIIMVEVITRRKPFDEIGGAPRIMAVANGTRPPLIKL 180
|||
Db 196 MADEVEGNSYSEKCVFSWGIIIMVEVITRRKPFDEIGGAPRIMAVANGTRPPLIKL 255
|||
Qy 181 PKPIESLMTRCWSKDPSPQSRPSMEIYKIMTHLMRYFPGADEPLQYPCQ 228
|||
Db 256 PKPIESLMTRCWSKDPSPQSRPSMEIYKIMTHLMRYFPGADEPLQYPCQ 303
|||

RESULT 3

US-09-757-982-5
/ Sequence 5, Application US/09757982
/ Patent No. US20020094559A1
/ GENERAL INFORMATION:
/ APPLICANT: Acton, Susan
/ TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
/ FILE REFERENCE: MNI-050
/ CURRENT APPLICATION NUMBER: US/09/757,982
/ PRIOR FILING DATE: 2001-01-10
/ PRIOR APPLICATION NUMBER: 09/163,115
/ PRIOR FILING DATE: 1998-09-29
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 5
/ LENGTH: 455
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-757-982-5

Query Match 29.6%; Score 371; DB 10; Length 455;

Best Local Similarity 37.7%; Pred. No. 7,4e-26;
Matches 81; Conservative 40; Mismatches 80; Indels 14; Gaps 8;

Qy 2 ELRLSRVNHPIVLYGACINP--VCLVMEYVAGSGSLVNVLHG--AEPLPYTTAAHMS 57
|||
Db 53 EAEILSVLSHRNIIFQYGVILEPPNYGIVTEYASLSGLDYINSNSEEM--DMDHIMT 109
|||
Qy 58 WCLOSGVAYLHSMQPKALIHRLDKPNLLVAGGTVLKICDFTACDICTHMTNNK-- 115
|||
Db 110 WATDVAKGMHYLHMEAVKVIHRLDKSRNVIAADG-VLKICDFG--ASRPHNHTHMSLV 167
|||
Qy 116 GSAAMMAPEVEGNSYSEKCVFSWGIIIMVEVITRRKPFDEIGGAPRIMAVANGTRP 174
|||
Db 168 GTFPMMAPEVIQSPVSECTDITYSGVLMEMLTREVPFGLGEG--LQVAMLVENGERL 225
|||
Qy 175 PLIKNLKPPIESLMTRCWSKDPSPQSRPSMEIYKIM 209
|||
Db 226 TIPSCEPSFAELHQCWEADAKKRPSPFQIISIL 260
|||

RESULT 4

US-09-862-027-19
/ Sequence 19, Application US/09862027
/ Patent No. US20020142428A1
/ GENERAL INFORMATION:
/ APPLICANT: Hodge, Martin R.
/ TITLE OF INVENTION: No. US20020142428A1 Kinases and Uses Thereof
/ FILE REFERENCE: 35800/234862
/ CURRENT APPLICATION NUMBER: US/09/862,027
/ CURRENT FILING DATE: 2001-05-21
/ PRIOR APPLICATION NUMBER: US 09/345,473
/ PRIOR FILING DATE: 1999-06-30
/ NUMBER OF SEQ ID NOS: 82
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 19
/ LENGTH: 394
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-862-027-19

Query Match 29.0%; Score 363.5; DB 10; Length 394;
Best Local Similarity 37.7%; Pred. No. 2,9e-25;
Matches 83; Conservative 36; Mismatches 84; Indels 17; Gaps 6;

Qy 2 ELRLSRVNHPIVLYGACINP--VCLVMEYVAGSGSLVNVLHGAEPPLPYTTAAHMSWC 59
|||
Db 50 EAKLFAMLKHPNIIALRGVCLKEPNLCLVMEFARGGPLNRVLSGKXIPPDI---LVNMA 105
|||
Qy 60 LQCSQGVAYLHSMQPKALIHRLDKPNLLV-----AGTVLKICDFTACD--IQTHM 111
|||
Db 106 VOIARGNVIHDEAIVPIIHRDLKSNITLILQKVENGDLSNKKILKITDFGLAREMHTTK 165
|||
Qy 112 TNNKSAAMMAPEVEGNSYSEKCVFSWGIIIMVEVITRRKPFDEIGGAPRIMAV--HN 170
|||
Db 166 MSAAGTYAMMAPEVIRASMSKSDVSYGVLLMELLTGEVPRRGIDG--LRVAYGVAMN 223
|||
Qy 171 GTRPPLIKNLKPPIESLMTRCWSKDPSPQSRPSMEIYKIMT 210
|||
Db 224 KLAIPISTCEPSPAKLMEDCWNPDPHSRPSFTIILDQLT 263
|||

RESULT 5

US-10-014-882-2
/ Sequence 2, Application US/10014882
/ Patent No. US20020107384A1
/ GENERAL INFORMATION:
/ APPLICANT: Hu, Yi
/ APPLICANT: Kieko, James
/ APPLICANT: Donoho, Gregory
/ TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding r
/ FILE REFERENCE: LEX-0279-USA
/ CURRENT APPLICATION NUMBER: US/10/014,882
/ CURRENT FILING DATE: 2001-12-11


```

Query Match Similarity 26.8%; Score 335.5; DB 10; Length 835;
Best Local Similarity 37.4%; Pred.No.2.1e-22;
Matches 80; Conservative 41; Mismatches 78; Indels 15; Gaps
8

OY 2 ELRQLSRVNHPIIVKLYGACL--NPVCLVMEYAEAGGSLYVNLHGAEPPLPYTAAHAMSW 58
D 509 EVSILLCQNLHPCVQFVPGACLDPSQFALVTVQYISGSGSLFSLH--EQRIIDLQSKLII 566
OY 59 CLQCSQGAVALHSM-QPKALIHRLDKPRPLLVAAGTVLKICDFGACDIQF---HMTN 113
D 567 ADVVAKGMEYLSLQF--IHRDLNSHILLIYEDGHA-VADFGSRPLQSLDEDMTK 623
OY 114 NKGSAAMAPAEVF-EGSNYSEKCDVFSWGIILMEVITRRKPRPEDEIGAPFRIMMAVHNGT 172
D 624 QGNLRMAAPAEVFTQCTRTIKADVFSYSLCMEWLLTGEIPFAHLKPAADAADMAVHH-I 682
OY 173 RPLIKNLKPIESLMTRCWSKDPSPORMEETV 206
D 683 RPIGYSIPKPISSLIRGMNACPEGRPFESEV 716

RESULT 9
US-09-771-161A-197
; Sequence 197, Application US/09771161A
; Patent No. US20020110811A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
; FILE REFERENCE: 802620-2005.1
; CURRENT APPLICATION NUMBER: US/09/771.161A
; CURRENT FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 197
; LENGTH: 966
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-771-161A-197

Query Match 26.5%; Score 331.5; DB 10; Length 966;
Best Local Similarity 34.3%; Pred.No.5.7e-22;
Matches 74; Conservative 36; Mismatches 89; Indels 17; Gaps
8

OY 2 ELRQLSRVNHPIIVKLYGACLNP--VCLVMEYAEAGGSLYVNLHGAEPPLPYTAAHAMSW 59
D 204 DIKHLRKLKHPRIIAFKGVCTQAPCCIMECANQGLYEVLRAKGI---IPRLLVDS 260
OY 60 LQCSQGAVALHSMQPKALIHRLDKPRPLLVAAGTVLKICDFGACDIQFHTNNK--GS 117
D 261 TGIAGSMYVLIHAK--IHRDLKSPN-VLVHTDAVKISDGTSEKSLSDKSTKMSFAGT 316
OY 118 AAMAPAEVFEGSNYSEKCDVFSWGIILMEVITRRKPRPEDEIGAPFRIMMAV-HNGRPL 176
D 317 VAMAPAEVIRNEPVSEKDWISFGVLMWELLTGEIPYKQVDSA--IIVGVSNSLHLPV 374
OY 177 IKNLKPIESLMTRCWSKDPSPORMEETV 212
D 375 PSTCPDGFKILMKQTKWQSKRNRPSEFRQ--ILMHL 407

RESULT 10
US-09-840-704-5
; Sequence 5, Application US/09840704
; Patent No. US20020122801A1
; GENERAL INFORMATION:
; APPLICANT: Dednar, Shoukat

```

```

APPLICANT: Hannigan, Greg
TITLE OF INVENTION: Intergryn-Linked Kinase and its Uses
FILE REFERENCE: KIN-200N
CURRENT APPLICATION NUMBER: US/09/840,704
CURRENT FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: 09/566,906
PRIOR FILING DATE: 2000-05-09
PRIOR APPLICATION NUMBER: US08/752,345
PRIOR FILING DATE: 1996-11-19
NUMBER OF SEQ ID NOS: 16
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 263
TYPE: PRT
ORGANISM: H. sapiens
FEATURE:
NAME/KEY: Other
LOCATION: (1)...(263)
US-09-840-704-5

Query Match      25.6%; Score 320; DB 10; Length 263;
Best Local Similarity 34.3%; Pred. No. 1,4e-21;
Matches 74; Conservative 43; Mismatches 91; Indels 8; Gaps 6;

Oy      2 ELRQSRVNHNNIVKLYACLNLP--VCLVMEYAEQGSILYNYLHGAEPLPYTAAHAMSMC 59
Db      51 EYALIMKRIRHPIYILFMGAVTQPPNLSIVTEYLSGSLYRLHLHSGAREQLDERRLSMA 110
Oy      60 LQCSQGVAYLHSMQPKALIHRLDKPNNLLVAGTVLTKICDFG-TACIOIOTHTNNK--G 116
Db      111 YDVAKGMNYLHNRNP-PIVHRDLKSPNLLVDKKYTV-KVCFGLSRLKASTFLSSKSAAG 168
Oy      117 SAAMAPVPEFGSNYSKCDYFSGCIIIMEVITRRKPPDEIGGPAFRIMAVAHNGTRPPL 176
Db      169 TPENWAPVPLVDEBSNEKSDYVSFGVILMELATLQQPWGNL-NPAQVVAAGVFKCKRLEI 227
Oy      177 IKNLEPKIESLMTRCWSDKPSQPSMEIYIMTHL 212
Db      228 PRNLNPQVAALIEGCVINEPWKRPSFATIMDLRPL 263

RESULT 11
US-09-904-389-2
Sequence 2, Application US/09904389
Patent No. US20020129404A1
GENERAL INFORMATION:
APPLICANT: Clendenen, Stephanie K.
TITLE OF INVENTION: CTRL HOMOLOGUE FROM MELON
FILE REFERENCE: 4257-0029.30
CURRENT APPLICATION NUMBER: US/09/904,389
CURRENT FILING DATE: 2001-07-12
PRIOR APPLICATION NUMBER: US 60/218,307
PRIOR FILING DATE: 2000-07-14
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 850
TYPE: PRT
ORGANISM: Cucumis melo
FEATURE:
NAME/KEY: VARIANT
LOCATION: (154)...(154)
OTHER INFORMATION: Xaa = Any, Amino Acid
US-09-904-389-2

Query Match      25.0%; Score 313.5; DB 10; Length 850;
Best Local Similarity 33.2%; Pred. No. 2e-20;
Matches 74; Conservative 47; Mismatches 89; Indels 13; Gaps 8;
Oy      2 ELRQSRVNHNNIVKLYACLNLP--VCLVMEYAEQGSILYNYLH--GAEPLEPYTTAAHAMS 57
Db      622 EYALIMKRIRHPIYILFMGAVTQPPNLSIVTEYLSGSLYRLHLHSGAVDI---DETRRIN 678

```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:58:31 ; Search time 3101 Seconds
(without alignments)

2139.774 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHPIVIXYGA.....MTHLMRYFFGADEPLQYPCQ 228

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% ,

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141754_4342/app_query.fasta_1.391
-DB=GenEmbl -DFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09830144 @CGN_1.1_3637/runat_04122002_141754_4342 -NCPUs=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: .gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1252	100.0	1704	6	E38399 NF-kappa B
2	1252	100.0	1705	9	AB009358 Homo sapi
3	1252	100.0	1745	9	AF218074 Homo sapi
4	1252	100.0	2443	10	D76446 Mouse mRNa
5	1252	100.0	2757	9	BC017715 Homo sapi
6	1252	100.0	2769	6	AX377912 Sequence
7	1252	100.0	2769	9	AB009356 Homo sapi
8	1252	100.0	2785	6	E38397 NF-kappa B
9	1252	100.0	2850	9	AB009357 Homo sapi
10	1252	100.0	2866	6	E38398 NF-kappa B
11	1252	100.0	3107	10	BC006665 Mus muscu
12	1221	97.5	2812	5	U92030 Xenopus lae
13	1133	90.9	195147	2	AC114407 Mus muscu
14	700	55.9	3349	3	AF199466 Drosophil
15	700	55.9	3386	3	AF051953 Drosophil
16	599	47.8	32564	2	AC014558 Drosophil
17	599	47.8	164942	3	AC011758 Drosophil
18	599	47.8	302303	3	AE003571 Drosophil
19	392.5	31.3	3138	9	HSMSTM
20	392.5	31.3	3454	6	AX337846 Sequence
21	392.5	31.3	3454	9	HARNMLK2
22	376	30.0	1445	5	AF265343 Danio rer
23	375.5	30.0	8125	3	AF481923 Dictyoste
24	374	29.9	1267	8	AY133876 Arabidops
25	374	29.9	1798	8	AY046026 Arabidops
26	373	29.8	1631	8	AY085535 Arabidops
27	373	29.8	1795	8	AY125513 Arabidops
28	372.5	29.8	3931	9	AF251442 Homo sapi
29	372	29.7	1429	10	AB049732 Mus muscu
30	372	29.7	3146	10	AB049731 Mus muscu
31	371	29.6	1365	6	AR119791 Sequence
32	371	29.6	1365	6	AR126751 Sequence
33	371	29.6	1365	6	AR128911 Sequence
34	371	29.6	1365	6	AR130842 Sequence
35	371	29.6	1365	6	AR138887 Sequence
36	371	29.6	1365	6	AR141355 Sequence
37	371	29.6	1368	9	AB049734 Homo sapi
38	371	29.6	1370	9	AF325454 Homo sapi
39	371	29.6	1571	9	BC001401 Homo sapi
40	371	29.6	2120	6	AR119790 Sequence
41	371	29.6	2120	6	AR126750 Sequence
42	371	29.6	2120	6	AR128910 Sequence
43	371	29.6	2120	6	AR130841 Sequence
44	371	29.6	2120	6	AR138886 Sequence
45	371	29.6	2120	6	AR141354 Sequence

ALIGNMENTS

RESULT 1

[illegible]

OY		121	MeLApProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp	140
Db		586	ATGGCACTGAAGTTTGTGAAGTAGTAATTACAGTGAAAAAATGTGACGTCCTCAGCTGG	645
OY		141	GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro	160
Db		646	GGTATTATCTTTGGAAAGTAATACCGGTCCGAAAACCCTTTGATGAGATTGTGGCCCCA	705
OY		161	AlAPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu	180
Db		706	GCTTTCCGAATCATGTGGCGCTTCATATAAGTACTGCACACCATGATTAATAAATTTA	765
OY		181	ProlYsProIleGluSerLeuMetThrArgCysTrpSerLeuYsAspProSerGlnArgPro	200
Db		766	CCTAAGCCCATTGAGAGCGCTGATGACTCGTGTGGTGCTTAAAGATCCTTCCCAAGGCCCT	825
OY		201	SerMetGluGluIleValLysIleMetThrHisLeuMetArgTrpPheProGluValAsp	220
Db		826	TCAATGGAGGAAATTGTGAATAATATGACTTCATCTGATCGCGTACTTCCAGAGCAGAT	885
OY		221	GluProLeuGlnTyrProCysGln	228
Db		886	GAGCCATTACAGTATCCTTGTCAG	909
RESULT 2				
AB009358				
LOCUS				
DEFINITION	Homo sapiens mRNA for TGF-beta activated kinase 1c, complete cds.	1705 bp	mRNA	linear PRI 20-JAN-2001
ACCESSION	AB009358			
VERSION	AB009358.2	GI:8978251		
KEYWORDS	TAK1c; TGF-beta activated kinase 1c.			
SOURCE	Homo sapiens cell_line:Hela CDNA to mRNA, clone:PT7TAK1c.			
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE				
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.			
TITLE	TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an NF-kappa B-inducing kinase-independent mechanism			
JOURNAL	Biochim. Biophys. Res. Commun. 243 (2), 545-549 (1998)			
MEDLINE	98153801			
REFERENCE				
AUTHORS	Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.			
TITLE	Alternative splicing and gene structure of the transforming growth factor beta-activated kinase 1			
JOURNAL	Biochim. Biophys. Acta 1517 (1), 46-52 (2000)			
MEDLINE	20568765			
REFERENCE				
AUTHORS	Sakurai,H., Shigemori,N., Hasegawa,K., Sugita,T. and Guesdon,F.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-DEC-1997) Francois Guesdon, University of Sheffield, Royal Hallamshire Hospital, Division of Molecular and Genetic Medicine, Functional Genomics Group, Glossop road, Sheffield S10 2JF, United Kingdom (E-mail:l.a.guesdon@sheffield.ac.uk,			
COMMENT FEATURES				
source	On Jul 8, 2000 this sequence version replaced gi:2924627.			
	location/Qualifiers			
	1..1705			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="pJT7TAK1c"			
	/cell_line="Hela"			
	1..1557			
	/note="TAK1c"			
	/codon_start=1			
	/product="TGF-beta activated kinase 1c"			
	/protein_id="BAZ5027.2"			
	/db_xref="gi:8978251"			
	/translation="MSTSAASSSSSSSAGSEMIEAPQVLNFEEDIDYKEIEVEVGVGR			
	GAFGVCAKMRKADVAIKQIESESERKAFIVEERQSIRVNHPVIIVKYGACLNPPVIC			
	VMEVREGSLNVLDHGAEPIPYTYDAARMSWCLCCSGSVAYILHMOPALIHRLDKPEF			
	NNLLVAGSTVTKICDPFTACDLCQHMTNNKSAAAMMAPEVREGSNYSERKCDVFWMGII			

LMEVITRRKPFDEIGPAPRIMVAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQRS
MEQVPMATNDTIKRLESKLLKNOAKQOSESGRLSLGSASRGVESLIPPTSEGRKMSADM
SEIARIAATTAYSKPKGHRKTAFSGNILDVPEIVISGNQPRRRSTODLTVTGTGP
GOVSRRSSPSVRMITTSGPTSEKPTRSHPWTPDDSTDTNSDINSIPWAYILTLHQLO
ARTSRTGPG"
511 a 381 c 402 g 411 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred. No.: 3.26e-117 Length: 1705
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x AB009358 (1-1705)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db 226 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 285
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyCysLeuTyrAsnVal 40
Db 286 TGCTTGAATCCAGTGTCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTG 345
QY 41 LeuHisGlyAlaGluProLeuTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 346 CTGCATGTGCTGAACCATTGCCATATTATACCTGCTGCCACGAATGAGTTGGTGTGTTA 405
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db 406 CAGTGITCCCAAGAGTGGCTTATCTTCACAGATGCAACCCAAACGGCTAATTCACAGG 465
QY 81 AspLeuIysProProIenLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db 466 GACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGGACAGTCTTAAAAATTTGTGAT 525
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTrp 120
Db 526 TTTGGGTACAGCCCTGTGACATTCAGACACACATACCAATAACAGGGGAGTGTGCTTGG 585
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluIysCysAspValPheSerTrp 140
Db 586 ATGCGACCTGAAGCTTTTGAAGGTAGTAAATACAGTGAAAAATGTGACGCTTCAGCTGG 645
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgGlyProPheAspGluIleGlyGlyPro 160
Db 646 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCCTTTTGTATGAGATTGTTGCCCA 705
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db 706 GCTTTCCGAATCATGTGGGCTGTTCATAATGGTACTCGACCCACCTGATAAAAAATTTA 765
QY 181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProSerGlnArgPro 200
Db 766 CCTAAGCCCATTTGAGACCTTGATGACTCGTTGTTGGTCTTAAAGATCCTTCCACGCGCCT 825
QY 201 SerMetGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 826 TCAATGAGGAAATTTGAAAAAATAGTACTCTTGATCGGTACTTTCCAGGAGCAGAT 885
QY 221 GluProLeuGlnTyrProCysGln 228
Db 886 GAGCCATTACAGTATCCTTGTCTAG 909
RESULT 3
AF218074 AF218074 1745 bp mRNA linear PRI 18-DEC-2000
LOCUS Homo sapiens TGF beta-activated kinase splice variant d (TAK1)
DEFINITION mRNA, complete cds.
ACCESSION AF218074

AF218074.1 GI:6746614
.
Homo sapiens.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1745)
Dempsey,C.E., Sakurai,H., Sugita,T. and Guesdon,F.
Alternative splicing and gene structure of the transforming growth
factor beta-activated kinase 1
Biochim Biophys. Acta 1517 (1), 46-52 (2000)
20568765
JOURNAL
MEDLINE
PUBMED
11118615
REFERENCE
2 (bases 1 to 1745)
Dempsey,C.E. and Guesdon,F.
Direct Submission
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..1745
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1..1745
/genes="TAK1"
CDS
7..1482
/genes="TAK1"
/notes="TAK1d"
/codon_start=1
/product="TGF beta-activated kinase splice variant d"
/protein_id="AAF27652.1"
/db_xref="GI:6746615"
translations="MSTAGAASSSSSSSSAGIEAPSOVLNFEIDYKEIEVEVVVGR
GAFVVCCKAKRAKDVAIKQIESESRKAFIVELROLSRVNHNIVKLYGACLNPVCL
VMEYAGGSLYNVLHGAELPIYTAAHAMSWCLQCQGVAYLHSMQPKALIHRDLKPP
NLLLVAGGTVLKICDFGTACDIQTHMNNKNSAAMWAPVEFGSNYSKCDVPSWGII
LMEVITRRKPFDEIGPAPRIMVAVHNGTRPPLIKNLKPKPIESLMTRCWSKDPQRS
MEQVPMATNDTIKRLESKLLKNOAKQOSESGRLSLGSASRGVESLIPPTSEGRKMSADM
MEQVPMATNDTIKRLESKLLKNOAKQOSESGRLSLGSASRGVESLIPPTSEGRKMSADM
SEIARIAATTAYSKPKGHRKTAFSGNILDVPEIVISGNQPRRRSTODLTVTGTGP
SEIARIAATTNGQPRRRSIODLTVTGTGPQVSSRSSSPSVMITTSPTSEKPTR
SHPWTPDDSTDTNSDINSIPWAYILTLHQLOARTSRTGPG"
BASE COUNT 526 a 375 c 412 g 432 t
ORIGIN
Alignment Scores:
Pred. No.: 3.35e-117 Length: 1745
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x AF218074 (1-1745)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db 232 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 291
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 292 TGCTTGAATCCAGTGTCTTGTGTGATGAATATGCTGAAGGGGGCTCTTTATATAATGTG 351
QY 41 LeuHisGlyAlaGluProLeuTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 352 CTGCATGTGCTGAACCATTCGCCATATTATACTGCTGCCACGAATGAGTTGGTGTGTTA 411
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db 412 CAGTGTTCCTCAAGAGTGGCTTATCTTCACAGATGCAACCCAAACGGCTAATTCACAGG 471
QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db 472 GACCTGAAACCAACCAACTTACTGCTGTTGTCAGGGGGACAGTCTTAAAAATTTGTGAT 531

QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIleGlySerAlaAlaATP 120
DB 532 TTTGGTACAGCCCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGCTGCTGG 591
QY 121 MetAlProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 592 ATGGCACCTGAACTTTTGAAGGTAAATTAACAGTGAAGAAATGACCTCTTCAGCTGG 651
QY 141 GlyIleIleuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 652 GGTATTATTCTTTGGAGAGTAAACGGCTGGAAACCTTTGATGAGATTGATGGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProleuIleLysAsnLeu 180
DB 712 GCTTTCCGATCATGTGGCTGTTCATTAATGTACTGCACCACTGATTAATAAATTTTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 772 CCTAACCCCTTACAGAGCTGATGACTCTGTTGTTGTTGTTAAAGATCTCTTCCAGGCCCT 831
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 832 TCAATGAGAGAAATGTGAAAATAATGACTCACTGATCGGGTACTTCCAGAGAGAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
DB 892 GAGCCATTACAGTATCTTGTCCAG 915
RESULT 4
MUSTAKI 2443 bp mRNA linear ROD 04-FEB-1999
LOCUS Mouse mRNA for TAK1 (TGF-beta-activated kinase), complete cds.
DEFINITION D76446.1 GI:1167505
ACCESSION D76446
VERSION TAK1; TGF-beta-activated kinase; protein kinase.
KEYWORDS Mus musculus cDNA to mRNA.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Yamaguchi, K., Shirakabe, K., Shibuya, H., Irie, K., Oishi, I., Yeno, N.,
Taniguchi, T., Nishida, E. and Matsumoto, K.
TITLE Identification of a member of the MAPKK family as a potential
mediator of TGF-beta signal transduction
JOURNAL Science 270 (5244), 2008-2011 (1995)
MEDLINE 96123277
REFERENCE 2 (bases 1 to 2443)
AUTHORS Matsumoto, K.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-1995) Kunihiko Matsumoto, Faculty of Science,
Nagoya University, Department of Molecular Biology, Furou-chou,
Chikusa-ku, Nagoya, Aichi 464-01, Japan
(E-mail: g44177a@nuc.cc.nagoya-u.ac.jp, Tel: 052-789-3000,
Fax: 052-789-3001)
FEATURES
source Location/Qualifiers
1..2443
/organism="Mus musculus"
/db_xref="taxon:10090"
157..1896
/product="TAK1 (TGF-beta-activated kinase)"
/protein_id="BA01184.1"
/db_xref="GI:1167506"
/translation="MSTASASSSSSSASEMIEAPQVLNFEEDYKEIEVEVGR
GAFGVCAKWRKADVAIKOIESSEKAFIVEIROLRSRVHNPVILVYGCALNPCL
VMEVAGSLVNVHAGAPLIPYTAAMMSVCLCCOSQVATLHMOPLALHRLKXP
NLLVAGSTVLKICDPGACDIQTHTMNNKSSAAMAEVVEGSSNYSKCVPEWGII
LMEVITRRKPPDEIGPAPRIMAVHNSTRPLIKNLPKPIESLMTKWSDPESQRS
MEEIVKIMTHLMRYFPQADELPQPCQYSDSGSNSATSGFMDIASNTNSKSDTN
MEOVPAINDTIKRLSKLKNQAKQSSGSLSGASGSVSESLPTSEKNSADM
SEIARIVATAGNQPRRSIODLIVTTEGQVSRSSPSVAMITTSGETSEKPAR
SHPTPDDSTDTNSDNSIPMAVILTLHQDPLAPCPNSKSMVAFDHCMAQEVYK

VOTETALLORKELEVALEDOBEKOQNTSRVLVOEHKKLLDENKSLSTYYOQCKOLE
VIRSQQKXQGS"
BASE COUNT 669 a 568 c 647 g 559 t
ORIGIN
Alignment Scores:
Pred. No.: 4 87e-117 Length: 2443
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-830-144-2_copy_76_303 (1-228) x MUSTAKI (1-2443)
QY 1 ValGluLeuArgGluLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 382 GTGGAGCTCCGGCAGTGTGGCCGTGTGAACCATCTTAACATTGTCAAGTTGTACGAGCC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 442 TGCCGTGATCCAGTATGTCTTGATGGAATATGACAGAGGGGCTCATTTGATTAATGTG 501
QY 41 LeuHisGlyIleGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 502 CTGCATGTGCTGAACCATTCCTTACTACTGCTGCTTATGCACTGACCTGCTTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 562 CAGTGTCCCAAGAGAGTGCGCTTACCTGCACAGCATGACGCCAAGCGGTGATTACAGG 621
QY 81 AspleuLysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 622 GACCTCAAGCCTCCMAACTTGCTGTGGTTGCAGAGGAGCACTGTTAATAAATCTCGCAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIleGlySerAlaAlaATP 120
DB 682 TTTGTTACAGCTTGTGTCATCCAAACACACATGACCAATATTAAGAGAGTCTGCTGG 741
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 742 ATGGCGCTGAAAGTGTGAAGGAGCAATTACAGTGAAGAAAGTGATGCTTCACTGG 801
QY 141 GlyIleIleuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
DB 802 GGTATTATCTCTGGGAGAGATTAACACGCCGGAACCTTCATGATGATCGGTGCCCCA 861
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProleuIleLysAsnLeu 180
DB 862 GCTTTCAGATCATGTGGCTGTTCATTAATGGCACTGCACCACTGATCAAAAATTTTA 921
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 922 CCTAAGCCATTGAGAGCTTGAAGTACACGCTGTGTTGGTTAAGGACCATCTCAGGCCCT 981
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 982 TCAATGAGAGAAATGTGAAAATAATGACTCACTGATGCGTACTTCCAGAGAGCGAT 1041
QY 221 GluProLeuGlnTyrProCysGln 228
DB 1042 GAGCCATTACAGTATCTTGTCCAG 1065
RESULT 5
LOCUS BC017715 2757 bp mRNA linear PRI 06-DEC-2001
DEFINITION Homo sapiens, mitogen-activated protein kinase kinase 7,
clone MGC:21263 IMAGE:3906857, mRNA, complete cds.
ACCESSION BC017715
VERSION BC017715.1 GI:17389342
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2757)
Strausberg, R.
Direct Submission
Submitted (03-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapps-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAK plate: 22 Row: i Column: 18
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507360.

FEATURES
source
1..2757
/organism="Homo sapiens"
/db_xref="LocusID:6885"
/db_xref="taxon:9606"
/clone="MGC:21263 IMAGE:3906837"
/issue_type="uterus, leiomyosarcoma"
/clone_lib="NIH MGC 71"
/lab_host="DH108"
/note="Vector: pCMV-SPORT6"
198..1937
/codon_start=1
/product="mitogen-activated protein kinase kinase kinase 7"
/protein_id="AAH17715.1"
/db_xref="GI:17389343"
/translation="NSTASASSSSSSAGEMIEAPSOVLPNEIDYKEIEVEEVVGRGAGVVKAKRAKDAIKQIESERKAFIVELQLSRVNHNPVILYGLACLNPVCLVMEYAGGSLYNVLHGAEPYPYTAHMSWCLQSGVAYLHNSQPKALLHRLKPPNULLVAGGVLKICDFGTACDIQTHNNKSGAAWAEVEPEGSYSKCDVFSWGIIIMEVITRRPFDEIGGPAFRIMWAVHNGTRPELIKNLPKPIESLMTKCSKDPQSRPSMBEIVKIMTHLMRYFPGADEPLQPCQSDQSNASTSGSFMDIASTNTSNKSDTNMEQVPATNDTIKRLSKLLKNOAKQSGSLGASRGSSVESLPPTSEGRKMSADMSEIPIARTTNGQPPRSIODLTVTGPEQVSSRSRSPSVRMITTSQPTSEKPTRSHPTWDDSTDNGSDNSTPMAYLFDLHQLPACPNKESMAVFEQHCWQAQBYMKVTEIALLQRKQELVAELDDQKQNTSRVUQBHKLLDENKSLSTYYQCKRQKLEVIRSQQRKQGTG"
BASE COUNT 819 a 566 c 647 g 725 t
ORIGIN

Alignment Scores:
Pred. No.: 5.57e-117 Length: 2757
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BC017715 (1-2757)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyVala 20
|||||
Db 423 GTAGAGCTTCGGCAGTATCCCGGTGAACCATCTTAATATTGTAAGCTTTATGAGCC 482
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
|||||

Db 483 TGCTTGAATCCAGTGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 542
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
|||||
Db 543 CTGATGGTGTGAACCATTTGCCATATTATCTGCTGCCACCGCAATGAGTTGGTGTTTA 602
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
|||||
Db 603 CAGTGTTCCTCCCAAGAGTGGCTTATCTTACAGCATGCAACCCAAAGCGCTTAATTCACAGG 662
Qy 81 AspleuLysProProAsnLeuLeuValAlaGlyClyThrValLeuLysIleCysAsp 100
|||||
Db 663 GACCTGAAACCCACCAAACTTACTGCTGGTTCAGGGGGGACAGTTCTTAAATAATTTGTGAT 722
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
|||||
Db 723 TTTGGTACAGCTGTGACATTCACACACACATGACCAATAACAAGGGGAGTGTCTGTGG 782
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
|||||
Db 783 ATGGCACCCTGAAGTTTTTGAAGGTAGTAATTACAGTGAAAAAATGTGACGTCTTCAGCTGG 842
Qy 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
|||||
Db 843 GGTATTATCTTTGGGAAGTGTATAACCGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 902
Qy 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
|||||
Db 903 GCTTTCGAATCATGTGGGCTGTTCAATAATGGTACTCGACCACCATGATAAAAAATTTA 962
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
|||||
Db 963 CCTAAGCCCATTTGAGAGCTGATGACTCGTTTGGTCTAAAGATCCTTCCACGCGCCT 1022
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
|||||
Db 1023 TCAATGGAGGAATTTGTAAAAAATAATGACTCACTTGATGCGGTACTTTCCAGGAGCAGAT 1082
Qy 221 GluProLeuGlnTyrProCysGln 228
|||||
Db 1083 GAGCCATTACAGTATCTCTTGTCAG 1106
RESULT 6
AX377912 AX377912 2769 bp DNA linear PAT 18-MAR-2002
LOCUS
DEFINITION Sequence 107 from Patent WO0212338.
ACCESSION AX377912
VERSION AX377912.1 GI:19573976
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gillen, C., Wetzel, I., Whendt, S., Weihe, E. and Schaefer, M. K.
TITLE Screening method
JOURNAL Patent: WO 0212338-A 107 14-FEB-2002;
Gruenthal GmbH (DE)
FEATURES
source Location/Qualifiers
1..2769
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 811 a 565 c 640 g 753 t
ORIGIN

Alignment Scores:
Pred. No.: 5.6e-117 Length: 2769
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-09-830-144-2_COPY_76_303 (1-228) x AX377912 (1-2769)
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleuTyrgIyala 20
      388 GTAGAGCTTCGGCAGATTATCCCGTGAACCAATCCATATATTGTAAGCTTATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrgAlaGluGlySerLeuTyraAsnVal 40
      448 TGGTTAAATCCAGTGTCTGTGTGATGAAATAGCTGAAGGGGGCTCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyrgTyrgThraAlaAlaHisAlaMetSerTrpCysLeu 60
      508 CTGCATAGTGGTGAACCATTTGACATTTATCTGCTGCCACGACATGATGTTGTTT 567
QY      61 GlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
      568 CAGGTGTCCCAAGAGTGGCTTATCTTACAGATGCAACCAAGCGCTTATTCACAGG 627
QY      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThraValLeuLysIleCysAsp 100
      628 GACCTTAACACCAACTTACTGCTGTGTGACGGGGGACGTTCTAAATAATTGTGAT 687
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaAlaTrp 120
      688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAGGGGAGTCTCTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTyrgSerGluLysCysAspValPheSerTrp 140
      748 ATGGCACCTGAAGTTTGAAGTAGTAATTACAGTGAATAATGACGCTTCACGTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThraArgLysProPheAspGluIleGlyLysPro 160
      808 GGTATTATTCTTTGGAGATGATTAAGCGTCGGAACCCCTTTGATGAGATTGTGGCCCA 867
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180
      868 GCTTCCGATCATGTGGCTGTTCATATGTAATGTAATCCGACCACTGATTAATAATTTA 927
QY      181 ProLysProIleGluSerLeuMetThraArgCysTrpSerLysAspProSerGlnArgPro 200
      928 CCTAAGCCCATGAGAGCCTGATGACTCGTTGTGTTAAAGATCTTCCACAGGCCCT 987
QY      201 SerMetGluGluIleValLysIleMetThraHisLeuMetArgTyrgPheProGlyAlaAsp 220
      988 TCAATGAGAGAAATGTGAATAATGACTCACTTGATGCGGTACTTCCAGGACAGAT 1047
QY      221 GluProLeuGlnTyrgProCysGln 228
      1048 GAGCATTATACAGATCTTGTCTCAG 1071
RESULT 7
AB009356          2769 bp      mRNA      linear      PRI 04-MAR-1998
LOCUS      Homo sapiens mRNA for TGF-beta activated kinase 1a, complete cds.
DEFINITION      AB009356
ACCESSION      AB009356
VERSION      AB009356.1 GI:2924623
KEYWORDS      TAK1a; TGF-beta activated kinase 1a.
SOURCE      Homo sapiens lung cDNA to mRNA, clone_lhb:lambda gt11
              clone:PBSTAK1a.
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (sites)
AUTHORS      Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE      TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
              NF-kappa B-inducing kinase-independent mechanism
JOURNAL      Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE      98153801
REFERENCE      2 (bases 1 to 2769)
AUTHORS      Sakurai,H.
TITLE      Direct Submision
JOURNAL      Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
              laboratory, Tanabe Seiyaku Co. Ltd., 16-89, Kashima-3-chome.

```

```

Yodogawa-ku, Osaka, Osaka 532, Japan (E-mail:hsakurai@tanabe.co.jp,
Tel:++81 6 300 2571, Fax:++81 6 300 2593)
FEATURES
    source
        1..2769
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="PBSTAK1a"
            /tissue_type="lung"
            /clone_lib="Lambda gt11"
            163..1902
            /note="TAK1a"
            /product="TGF-beta activated kinase 1a"
            /protein_id="BA25025.1"
            /db_xref="GI:2924624"
            /translation="MSTASASSSSSSAGEMIEAPQVLFNEEIDYEKEIEVEEYVGR
            GAFVYCKAKMPAKDVALKOISESEERKAFIYELKOLSVNHPNIVKLYGACLNIVCL
            VMEYAEAGSGSLVNVLHGAEPYPYTAHAMSMCLQCSQGVAYHSMOPRALIHRLDKPP
            NLLVAGGVTLKICDPGTACDIQTMTNKKGSAMMAPEVEGSGSYSEKCDVFSWGITI
            LMEVITRRKPFDEIGEPARIMAVHNGTRPPLIKLKPPIESLMTKWSKDPQSPS
            MEIVAKIMTHMKYEPGADEPLQYPCQYSDQGSATSTGSPMDIATNTSNKSDTN
            SEIEARIAATTGNGOPRRSIQDLIVTGTPEQVSRSSPSVPMITTSPTSEKPTR
            SHPMPEDDSTDTNGSPNSIPMAVLTLDHQLAPCPNSKESMAVFEQHCMAQCYMK
            VQTEIALLLQKQELVAELDQDEBKQQTNSRLVQHKULDENKSLSTYYQCKKQLE
            VTRSQQKQKQGS"
BASE COUNT      811 a      565 c      640 g      753 t
ORIGIN
Alignment Scores:
Pred. No.:      5,6e-117      Length:      2769
Score:          1282.00      Matches:      228
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              9      Gaps:      0
US-09-830-144-2_COPY_76_303 (1-228) x AB009356 (1-2769)
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleuTyrgIyala 20
      388 GTAGAGCTTCGGCAGATTATCCCGTGAACCAATCCATATATTGTAAGCTTATGAGCC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrgAlaGluGlySerLeuTyraAsnVal 40
      448 TGGTTAAATCCAGTGTCTGTGTGATGAAATAGCTGAAGGGGGCTCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyrgTyrgThraAlaAlaHisAlaMetSerTrpCysLeu 60
      508 CTGCATAGTGGTGAACCATTTGACATTTATCTGCTGCCACGACATGATGTTGTTT 567
QY      61 GlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
      568 CAGGTGTCCCAAGAGTGGCTTATCTTACAGATGCAACCAAGCGCTTATTCACAGG 627
QY      81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThraValLeuLysIleCysAsp 100
      628 GACCTTAACACCAACTTACTGCTGTGTGACGGGGGACGTTCTAAATAATTGTGAT 687
QY      101 PheGlyThraAlaCysAspIleGlnThraHisMetThraAsnLysGlySerAlaAlaTrp 120
      688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAGGGGAGTCTCTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTyrgSerGluLysCysAspValPheSerTrp 140
      748 ATGGCACCTGAATTTTGAAGTACTTATTCAGTGAATAATGTAAGCTTCTCAGCTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThraArgLysProPheAspGluIleGlyLysPro 160
      808 GGTATTATTCTTTGGAGATTAAGCGTCGGAACCCCTTGATGAGATGATGTGGCCCA 867
QY      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleLysAsnLeu 180

```

```

Db 868 GCTTCCGAATCATGTGGCGCTTTCATAATGGTACTCGACACACACTGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCATTGAGAGCCTGATGACTCGTGTGGTCTTAAGATCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 988 TCAATGGAGGAAATGTGAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCATTAAGTATCCTTGTGCAG 1071

RESULT 8
E38397
LOCUS E38397 2785 bp DNA linear PAT 31-JAN-2002
DEFINITION NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same.
ACCESSION E38397
VERSION E38397.1 GI:18626977
KEYWORDS JP 2000197500-A/3.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2785)
AUTHORS Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLE NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
JOURNAL Patent: JP 2000197500-A 3 18-JUL-2000;
TANABE SEIYAKU CO LTD
OS Unidentified
PN JP 2000197500-A/3
PD 18-JUL-2000
PF 04-FEB-1999 JP 1999026803
PR
PI NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA, PI HIROSHI
HASEGAWA
PC C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99, PC
C12Q1/02.
PC G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,(C12N15/09,
PC C12R1/91),
PC C12N5/00,C12N15/00,(C12N15/00,C12R1/91)
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT source 1..2785
FT Location/Qualifiers
1..2785
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 827 a 565 c 640 g 753 t
ORIGIN

Alignment Scores:
Pred. No.: 5.63e-117 Length: 2785
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x E38397 (1-2785)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db 388 GTAGAGCTTCGGCAGTATCCCGTGTGAACCATCCTAATATTGAAAGCTTTATGGAGCC 447
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db 448 TGCATTGAATCAGTGTCTTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATAATGTG 507

```

```

Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
Db 508 CTGCATGTGTGCTGAACCATTCGCATATTACTGTGCTGCCAGCAATGAGTGGTGTTTA 567
Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 568 CAGTGTTCCTCCAGGAGTGGCTTATCTTCACAGCATGCAACCAAGCGCTAATTACACAGG 627
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 628 GACCTGAAACACCAACTTACTGCTGTTGAGGGGGGAGAGTTCTAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
Db 748 ATGGCACCTGAAGTTTGTGAAGGTAGTAATTACAGTGAATAATGTGACGTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATTCTTTGGGAAGTAGTAACGCGTCGAAACCCCTTTGATGAGATTGGTGGCCCA 867
Qy 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 868 GCTTTCCGAATCATGTGGCTGTTTCATAATGTGACTCCACCACCATGATAAAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTyrPheLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATTTGAGAGCCTGATGACTCGTGTGGTCTAAAGATCTCTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 988 TCAATGGAGGAAATGTGAAATAATGACTCACTTGATGGGTACTTTCCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTyrProCysGln 228
Db 1048 GAGCCATTACAGTATCCTTGTGCAG 1071

RESULT 9
AB009357
LOCUS AB009357 2850 bp mRNA linear PRI 04-MAR-1998
DEFINITION Homo sapiens mRNA for TGF-beta activated kinase 1b, complete cds.
ACCESSION AB009357
VERSION AB009357.1 GI:2924625
KEYWORDS TAK1b; TGF-beta activated kinase 1b.
SOURCE Homo sapiens lung cDNA to mRNA, clone_lib:lambda gt11
clone:pBSTAK1b.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE 1
AUTHORS Sakurai,H., Shigemori,N., Hasegawa,K. and Sugita,T.
TITLE TGF-beta-activated kinase 1 stimulates NF-kappa B activation by an
NF-kappa B-inducing kinase-independent mechanism
JOURNAL Biochem. Biophys. Res. Commun. 243 (2), 545-549 (1998)
MEDLINE 98153801
REFERENCE 2 (bases 1 to 2850)
AUTHORS Sakurai,H.
TITLE Direct Submission
JOURNAL Submitted (01-DEC-1997) Hiroaki Sakurai, Lead Generation research
Laboratory, Tanabe Seiyaku Co. Ltd.; 16-89, Kashima-3-chome,
Yodogawa-ku, Osaka, Osaka 532, Japan [E-mail:hsakurai@tanabe.co.jp,
Tel:++81 6 300 2571, Fax:++81 6 300 2593]
1..2850
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pBSTAK1b"
/tissue_type="lung"
/clone_lib="Lambda gt11"
163..1983
FEATURES
source
1..2850

```

CDS


```
/note="TAK1b"
/codon_start=1
/product="TGF-beta activated kinase 1b"
/protein_id="BAA25026.1"
/db_xref="GI:2924626"
/translation="MSTASASSSSSSAGEMIEAPSOVLNFEIDYKEIEVEVGR
GAFGVCAKKRAVDVAIKOIESESERAFVIEIROLSRNHPITVLGYACINPVL
VMEVAREGSLYNVLHGAPLPYTRTAHMSVCLDSCSOVALTHMOBKALIHDLKP
NLLVAGSTVLKICDFGTACDIQTHMTNKSAAWMAVEVEGNSISEKCDVPSGII
LMEVITRRKPPDEIGGPAFRIMAVHNGTRPLIKLPEKIESLMTKWSKDSQPS
MEIIVKMTLMRYPGADELPQCYSDSGNSATISGFMSTASTMSKSDTN
MEOVATNDTTIKRLSKLKNQAKOOSGSLSGASRGVSELPPTSGKMSADM
SEIRIATATTAYSKPKRGHRTAKTSFQNIILVPEIVISNGQRRRIODIYVGTPE
GVSSRSSPSVRMTTTSSTGPTSEKTRSHPTPPDSTDNDSIMAYITLDHQO
PLAPCNSKSNMAYEHOCHKMAQETMKVQITALLQKQELVALLEODEKQONTSR
LVEHKIKLDENKSLSTYYQCKQKQLEVIRSQOKROGTS"
```

```
BASE COUNT      832 a      588 c      656 g      774 t
ORIGIN
Alignment Scores:
Pred. No.:      5,786-117      Length:      2850
Score:      1252.00      Matches:      228
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      Gaps:      0
```

US-09-830-144-2_COPY_76_303 (1-228) x AB009357 (1-2850)

```
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrgIyAla 20
Db      388 GTAAGACCTTGAGAGTTATCCCGTGAACCATCTTAATTTGTAAGCTTTATGAGACC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrgIyAlaGluGlySerLeuTyrgAsnVal 40
Db      448 TGGTTGAATCCAGTGTCTGTTGATGAGATATGCTGAAGGGGCTCTTATATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyrgTyrgAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGGAACCATTTGCTTACTGCTGCCACGCAATGATGGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db      568 CAGGTGTTCCCAAGAGAGTGGCTTATCTTCAAGATGCCAACCCAAAGCGTTATTCACAG 627
QY      81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db      628 GACCTGAACCAACCAACTTACTGCTGTTGCAGGGGAGCAGTTCTAATAAATTTGTGAT 687
QY      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerIleAlaIleTrp 120
Db      688 TTTGGTACAGCCCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTCTCTTGG 747
QY      121 MetAlaProGluValPheGluGlySerAsnTyrgSerGluIysCysAspValPheSerTrp 140
Db      748 ATGGCACCTGAACTTTTGAAGTACTTAATTACAGTGAATAATGTGACCTTTCAGCTGG 807
QY      141 GlyIleIleLeuTrpGluValIleThrArgArgIysProPheAspGluIleGlyIlePro 160
Db      808 CGATTAATCTTTGGAGATGATACGCGTGGAAACCTTTGATGAGATTGTGTGGCCCA 867
QY      161 AlaPheArgIleMetSerTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db      868 GCTTTCGGAATCATGTGGGCTGTTCATAATGTGACTCGAACCACTGATTAATAAATTTTA 927
QY      181 ProIysProIleGluSerLeuMetThrArgCysTrpSerIysAspProSerGlnArgPro 200
Db      928 CTTAAACCCATTAGAGCCGTGATGACTCGTTGTGTTCTTAAGATCTTTCCAGGGCCCT 987
QY      201 SerMetGluGluIleValIysIleMetSerThrHisLeuMetArgTyrgPheProGlyAlaAsp 220
Db      988 TCAATGAGAGAAATTTGTAATAATATGACTCTGATGCGGTACTTTCCAGAGAGAGAT 1047
```

```
QY      221 GluProLeuGlnTyrgProCysGln 228
Db      1048 GAGCCATTACAGTATCTTGTCTCAG 1071
```

```
RESULT 10
E38398
LOCUS      E38398      2866 bp      DNA      linear      PAT 31-JAN-2002
DEFINITION      NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same.
ACCESSION      E38398
VERSION      E38398.1 GI:18626978
KEYWORDS      JP 2000197500-A/4.
SOURCE      unidentified.
ORGANISM      unidentified.
REFERENCE      1 (bases 1 to 2866)
AUTHORS      Sugita,N., Sakurai,H., Kageyama,N. and Hasegawa,H.
TITLES      NF-kappa B activation inhibitory drug targeting TAK1 and method for
identifying the same
PATENT      JP 2000197500-A 4 18-JUL-2000;
JOURNAL      TANABE SEIYAKU CO LTD
COMMENT      OS      Unidentified
PN      JP 2000197500-A/4
PD      18-JUL-2000
PF      04-FEB-1999 JP 1999026803
PR      NAOHISA SUGITA,HIROAKI SAKURAI,NORIKO KAGEYAMA,PI HIROSHI
HASEGAWA
PC      C12Q1/48,A61K31/00,A61K31/00,A61K45/00,C12N5/10,C12N9/99,PC
C12Q1/02,
PC      G01N33/15,G01N33/50,G01N33/566//C12N15/09,C12Q1/68,C12N15/09,
PC      C12R1.91),
PC      C12N5/00,C12N15/00,C12N15/00,C12R1.91)
CC      Strandedness: Double;
CC      Topology: Linear;
FH      Key      Location/Qualifiers
FT      source      1..2866      /organism='Unidentified'
```

```
BASE COUNT      848 a      588 c      656 g      774 t
ORIGIN
Alignment Scores:
Pred. No.:      5,826-117      Length:      2866
Score:      1252.00      Matches:      228
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      Gaps:      0
```

US-09-830-144-2_COPY_76_303 (1-228) x E38398 (1-2866)

```
QY      1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrgIyAla 20
Db      388 GTAAGACCTTGAGAGTTATCCCGTGAACCATCTTAATTTGTAAGCTTTATGAGACC 447
QY      21 CysLeuAsnProValCysLeuValMetGluTyrgIyAlaGluGlySerLeuTyrgAsnVal 40
Db      448 TGGTTGAATCCAGTGTCTTGTGATGAGATATGCTGAAGGGGCTCTTATATATG 507
QY      41 LeuHisGlyAlaGluProLeuProTyrgTyrgAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGGAACCATTTGCTTACTGCTGCCACCAATGATGTGTTA 567
QY      61 GlnCysSerGlnGlyValAlaTyrgLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
Db      568 CAGGTGTTCCCAAGAGAGTGGCTTATCTTCAAGCATGCAACCCAAAGCGTTAATTCACAG 627
QY      81 AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
```



```

Db 628 GACCTGAAACACCAAACTTACTGCTGCTGGTTGCAGGGGGGACAGTCTTAAAAAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120
Db 688 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACCAAGGGAGTGTGCTTGG 747
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrrp 140
Db 748 ATGCACACTGAAGTTTGTGAAGGTAGTAATACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy 141 GlyIleIleLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 808 GGTATTATTCCTTGGGAAGTGATAACCGTCGGAAACCCCTTGTATGAGATGTGTGGCCCA 867
Qy 161 AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 868 GCTTTCCGAATCATGTGGCTGTTTCATAATGGTACTCGACACACCACTGATATAAAATTTA 927
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrrpSerLysAspProSerGlnArgPro 200
Db 928 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTTGGTCTAAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTrrpPheProGlyAlaAsp 220
Db 988 TCAATGAGAGAAATTTGAAAAATAATGACTCACTGATGCGGTACTTTCAGGAGCAGAT 1047
Qy 221 GluProLeuGlnTrrpProCysGln 228
Db 1048 GAGCATATACGATATCCTTGTTCAG 1071

RESULT 11
BC006665
LOCUS Mus musculus, clone MGC:5989 IMAGE:3499247, mRNA, ROD 07-AUG-2002
ACCESSION BC006665
VERSION BC006665.1 GI:13879375
KEYWORDS MGC.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (27-APR-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunarathne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LIML at: http://image.llnl.gov
Series: IRAK Plate: 9 Row: h Column: 9
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, GenomeScan gene prediction, Similarity but not identity
to protein.

FEATURES
source
1. .3107
/organism="Mus musculus"

```

```

/db_xref="taxon:10090"
/map="EVB/N"
/clones="MGC:5989 IMAGE:3499247"
/tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
ductal Carcinoma, 5 month old virgin mouse."
/clone_lib="NCI_CGAP Mam6"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
148..1968
/codon_start=1
/product="Unknown (protein for MGC:5989)"
/protein_id="AAH06665.1"
/db_xref="GI:13879376"
/translacion="MSTASASSSSSSASEMIEAPSOVLNFBEBIDYKEIEVEVVGGR
GAFGVCKAKWRADVAIKOIESESRKAFIVELROLRYVNHPIVVKLYGACLNVPVCL
VMYAEAGSILYNVLHGAEPYPTAAHMSNCLQCSQGVAYLHSMOPKALIHRLKPP
NLVVVAGTGVLKICDFTGACDIQTHMTNNKSAWMAPEVEGSENTERSEKCDVFSWGII
LWEVITRRKPFDEIGGPAFRIMWAVHNGTRPILKNLPKPIESLMTRCWSKDPSPRS
MEEVPATNDTIKRLESKLNKQAQSEGRSLSIGASRGSSVESLPPTSSEGRMSADM
SEIBARIVTAAYSKPKRHKRTASFGNILDVPEIVISGNQPRRSIQDLTVTGTPEP
GOVSSRSRSPSRMITTSGPTSEKPARSHPEWDDSDTDNGSDNSIPMAVILTDHQLQ
PLAPCNKSKSMVFBQHKWAOEYIKVQTEIALLQKQELVAELDDQDERDQONTSR
LVQEHKKLLDENKSLSTYYOQCKQKQEVIRSQOQKRGTS"
BASE COUNT 898 a 687 c 772 g 750 t
ORIGIN
Alignment Scores:
Pred. No.: 6,36e-117 Length: 3107
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Dbs: 10 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x BC006665 (1-3107)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTrrGlyAla 20
Db 373 GTGGAGCTCCGGCAGTGTCCGCTGTGAACCATCTTACATTGTCAGTTGACGAGCC 432
Qy 21 CysLeuAsnProValCysLeuValMetGluTrrAlaGluGlyGlySerLeuTrrAsnVal 40
Db 433 TGCCTGAATCCAGTATGCTTGTGATGGAATATGACAGGGGGGCTCATTTATAATGTG 492
Qy 41 LeuHisGlyAlaGluProLeuProTrrTrrThrAlaHisAlaMetSerTrrCysLeu 60
Db 493 CTGCATGTGTCTGAACCATTCCTTACTACACTGCTGCTCATGCATGAGCTGGTGTTA 552
Qy 61 GlnCysSerGlnGlyValAlaTrrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db 553 CAGTGTTCACAGGAGTGGCTTACTGACAGCATGACGCCAACGCGCTGATTCACAGG 612
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 613 GACCTCAGCCTCCAAACTTCTGCTGTTCCAGAGGAGCAGTCTTAAAAAATCTGCCAT 672
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrrp 120
Db 673 TTTGGTACAGCTTGTGACATCCAAACACACATGACCAATAATAAAGGAGTGTGCTTGG 732
Qy 121 MetAlaProGluValPheGluGlySerAsnTrrSerGluLysCysAspValPheSerTrrp 140
Db 733 ATGGCGCCTGAAGTGTTTGAAGGTAGTCAATTACAGTAAAAAGTGTGATGCTTCAGCTGG 792
Qy 141 GlyIleIleLeuTrrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 793 GGTATTATTCCTCTGGGAAGTGATAACACGCCCGGAAACCCCTTCGATGAGATCGGTGGCCA 852
Qy 161 AlaPheArgIleMetTrrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 853 GCTTTCAGAAATCATGTGGGCTGTTTCATAATGGCACTGCACCACTGATCAAAAATTTA 912

```

QY 181 ProlysProIIeGIuseRleuMetThrArgCysTrpSerIysAspProSerGlnAnpPro 200
|
|
|
Db 913 CCTAACCCATTGAGAGCTTGATGACACGCTGTTGCTTAAGACCACTCAGCGCCT 972
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluYalaAsp 220
|
|
|
Db 973 TCATGTGAGGAAATTTGTGAAAAATAGACTCTGATCGGTACTTCCAGAGCGGAT 1032
QY 221 GluProLeuGlnTyrProCysGln 228
|
|
|
Db 1033 GAGCCATTACAGTATCCTGTCTCAG 1056
RESULT 12
XLU92030 2812 bp mRNA linear VRT 18-APR-1998
LOCUS Xenopus laevis TGF-beta-activated kinase TAK1 mRNA, complete cds.
DEFINITION U92030
ACCESSION U92030.1 GI:3057035
VERSION
KEYWORDS
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 2812)
Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
Matsumoto,K., Nishida,E. and Ueno,N.
Role of TAK1 and TAB1 in BMP signalling in early Xenopus development
EMBO J. 17 (4), 1019-1028 (1998)
JOURNAL 98130593
PUBMED 9463380
REFERENCE 2 (bases 1 to 2812)
Shibuya,H.
Direct Submission
Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES
source
1..2812
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
341..2191
/note="TGF-beta-activated kinase"
/codon_start=1
/product="TAK1"
/protein_id="AAC14008.1"
/db_xref="GI:3057036"
/translation="MSATSAMETETTPVLNFEIIDEYKEIEVEVVGRTFGVCKAKM
RGKDAIKOIESESRKAFIVELROLRSVNHPTIKLAGCLNPVCLIMEVAEGSLY
NVLHGAEPPIPYTAHAMSWICQAGYALHSMKPKALHRLDKPRLLIVAGSTYL
KICDFGTACDIDTHTNKGSAAMAAPVEFBSNYSKCDVFSKIIIMEVTRKRP
DEIGAPAFIRIMAVHNGTRPPLIKNLPRFISLMTKWSKDPORPSMEELVKIMTL
MQYPPGADVSLQPCQYSDGQSNSTGSCIDITSTNTNKSIDHIEPDQASAS
NDTIKRIEKLAKOHLKNOAKGTGSESLSPRSRSVESSEIRGRSPSLGSEK
RMSADMSLEARIASATAVVKRGRKRTASGNLIDVPEIIITAGNQQRRRVODL
TIVGVSSQESKNSRSSSPSRMTTTSQTPPRDKPRRLPAPPESSSTNSDSITK
AVLTDHQLPLAPCPNSKESMAVFBQCKPAQEWKQVITIALLLQKQELIELDQ
DEKDOQNSRLVQEHKHLIDENKSLSTYYQCKQLVIRSQQKRGQTS"

QY 21 CysLeuAnpProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAnpVal 40
|
|
|
Db 593 TGCCATAATCCGGATATGTTATGATGAGATGACGAAGAGGTTCTATGTAATGTT 652
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
|
|
|
Db 653 TTGCATGAGACTGAACCGCTGCTTACTATACAGCTGCTCATGCAATGAGTTGTTT 712
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
|
|
|
Db 713 CAATGTCCCAAGAGAGTTCATATTTATCATGACATGAAGCCGAAGCTTAATTACAG 772
QY 81 AspLeuLysProProAnpLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
|
|
|
Db 773 GACCTCAACACCAACCAACTTGCTGCTGGTGGAGGACCCGTTCTTAAGATTTGTGAC 832
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAspAnpLysGlySerAlaAlaTyr 120
|
|
|
Db 833 TTGCTACAGCCTTGATATTTACAGACTCACATGACTATATACAAAGAGTACGATGG 892
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
|
|
|
Db 893 ATGGCTCCAGAGATTGTTGAAGTAGCAATTACAGGAGAAAGTGACGATTTAGTTGG 952
QY 141 GlyIleIleLeuTyrGluValIleMetThrArgTyrProPheAspGluIleGlyIlePro 160
|
|
|
Db 953 GGAATTAATCTTTGGAGAGTAATACCCGAAGAAACCTTTCATATAAATGGTGTGCCA 1012
QY 161 AlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuIleLysAnpLeu 180
|
|
|
Db 1013 GCGTTCGGTATCATGTGTGGCTGTTCCAAATGTACTCGCCACATTAATTAATAATTG 1072
QY 181 ProlysProIIeGIuseRleuMetThrArgCysTrpSerIysAspProSerGlnAnpPro 200
|
|
|
Db 1073 CCTAAGCCTATTGAAGCTTAATGACTCGCTGCTGGTCCAAAGACTCCCCACAAAGCCT 1132
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluYalaAsp 220
|
|
|
Db 1133 TCATGTGAGGAGGATGTGCAAGATATGATCATCATGATGATTTTCTGTGAGCGGAT 1192
QY 221 GluProLeuGlnTyrProCysGln 228
|
|
|
Db 1193 GTTTCCTTACAGTATCCTGTCTCAG 1216
RESULT 13
AC114407 135147 bp DNA linear HTG 26-JUN-2002
LOCUS AC114407
DEFINITION Mus musculus clone RP23-51G1, WORKING DRAFT SEQUENCE, 5 ordered
pieces.
AC114407.3 GI:21592111
VERSION HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 135147)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-51G1
Unpublished
2 (bases 1 to 135147)
REFERENCE AUTHORS
TITLE
JOURNAL
1 (bases 1 to 135147)
2 (bases 1 to 135147)
Birren,B., Linton,L., Nusbaum,C., Lander,E., All,A., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,
Bouhgalter,B., Brown,A., Camarata,J., Campoliano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., Dehrelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagoe,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Labocque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,
Macdonald,P., Major,D., Marquis,N., Matthews,C., McCarthy,M.,

McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (08-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 135147)

REFERENCE

AUTHORS

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavskiy, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., LaRocque, K., Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jun 26, 2002 this sequence version replaced gi:21536031.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L23267

Center clone name: 51_G_1

----- Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 133686 bases at least Q40

Consensus quality: 134318 bases at least Q30

Consensus quality: 134777 bases at least Q20

Insert size: 151000; agarose-fp

Insert size: 134747; sum-of-contigs

Quality coverage: 11.0 in Q20 bases; agarose-fp

Quality coverage: 12.3 in Q20 bases; sum-of-contigs

- * NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. Gaps between the contigs
- * are represented as runs of N. The order of the pieces
- * is believed to be correct as given, however the sizes
- * of the gaps between them are based on estimates that have
- * provided by the submittor.
- * This sequence will be replaced
- * by the finished sequence as soon as it is available and
- * the accession number will be preserved.

* 1 12654: contig of 12654 bp in length
 * 12655 12754: gap of 100 bp
 * 12755 13487: contig of 733 bp in length
 * 13488 13587: gap of 100 bp
 * 13588 14337: contig of 750 bp in length
 * 14338 14437: gap of 100 bp
 * 14438 63285: contig of 48848 bp in length
 * 63286 63385: gap of 100 bp
 * 63386 135147: contig of 71762 bp in length.

FEATURES

source

1..135147
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone_lib="RP23-51G1"
 /Clone lib="RP23-51G1-23 Female Mouse BAC"

misc_feature

1..12654
 /notes="assembly_fragment"
 clone_end:SP6
 vector_side:left
 12755..13487
 /notes="assembly_fragment"
 13588..14337
 /notes="assembly_fragment"
 14438..63285
 /notes="assembly_fragment"
 63386..135147
 /notes="assembly_fragment"
 clone_end:T7
 vector_side:right

BASE COUNT 43081 a 25796 c 26087 g 39780 t 403 others
 ORIGIN

Alignment Scores:
 Pred. No.: 4..79e-103 Length: 135147
 Score: 1133.00 Matches: 212
 Percent Similarity: 93.86% Conservative: 2
 Best Local Similarity: 92.98% Mismatches: 8
 Query Match: 90.50% Indels: 7
 DB: 2 Gaps: 2

US-09-830-144-2_COPY_76_303 (1-228) x AC114407 (1-135147)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
 Db 22303 GTGAGAGTCGCGAGTGTTCACGTGTGAACCATCTTAATATTGATAAGTTGTATGGAGCC 22362
 Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
 Db 22363 TGCCTGCATCCAGTATGTCTTGTGATGGAATATACAGAGGGGGGCTCATTTGATAATGTG 22422
 Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
 Db 22423 CTGCATGTGTGTGAACCATTCGCTTACTACACTGCTGCTCATGCAGTCAGCTGGTGTTA 22482
 Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
 Db 22483 CAGCT-----TACCTGCACAGCATGCAGCCAAAGCTCTAATTACACAGG 22526
 Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
 Db 22527 GACCTCAAGCCTCCAAACTTCTGCTGTTGCAGAGGAGCAGTTCTAAAAAATCTGGAT 22586
 Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTyr 120
 Db 22587 TTTGATACAGCTTGTGACATCCAGACACACATGACCAATAATAAAGGAGTGTCTGTGG 22646
 Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTyr 140
 Db 22647 ATGGTCCCCGAAGTGTTTGAAGGTAGCAATTTACAGTGAAGAGTGTGATGCTCTCAGCTGG 22706
 Qy 141 GlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
 Db 22707 GGTATTATCTCTGGGAAGTGATAACACGCTGGAACCC---GATGAGATCGTGGCCCA 22763

QY 161 AlaPhaArg11LeuMetTrpAlaValHisAsnGlyThrArgProProLeuLeuLeuLeuLeu 180
Db 22764 GCCTTCAGATTCAATGCGACTGTTCAATATGCGACCTCGACCACTGATCAAAATTTA 22823
QY 181 ProLeuPro11LeuSerLeuMetThrArgCysTrpSerLeuAspProSerGlnArgPro 200
Db 22824 CTTACGCCCATTTGAGAGCTTGATGACACACTGTTGGTTTAAAGGATCTTCTCAAGCCCT 22883
QY 201 SerMetGluGluLeuValLeuValHisLeuMetTrpHisLeuMetArgTrpPheProGluAlaAsp 220
Db 22884 TCATATGAGAAATTTGTGAATAATAGACTCACTTGATGCGGATCTCCACAGAGGAT 22943
QY 221 GluProLeuGluTrpProCysGln 228
Db 22944 GAGCCATTACAGTATCCTTGACG 22967
RESULT 14
AF199466 3349 bp mRNA linear INV 17-APR-2000
LOCUS Drosophila melanogaster TGF-beta activated-kinase 1 homolog mRNA,
DEFINITION complete cds.
ACCESSION AF199466
VERSION AF199466
KEYWORDS AF199466.1 GI:6289098
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pelecyota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 3349)
AUTHORS Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M. C., Matsumoto, K.,
O'Connor, M. B., Shibuya, H. and Ueno, N.
TITLE TAK1 participates in c-Jun N-terminal kinase signaling during
Drosophila development
JOURNAL Mol. Cell. Biol. 20 (9), 3015-3026 (2000)
PUBMED 20221548
JOURNAL 10757786
MEDLINE 2 (bases 1 to 3349)
REFERENCE Takatsu, Y., Nakamura, M., Stapleton, M., Danos, M., Matsumoto, M.,
O'Connor, M. B., Shibuya, H. and Ueno, N.
AUTHORS Direct Submission
TITLE Submitted (28-OCT-1999) Developmental Biology, National Institute
JOURNAL for Basic Biology, 38 Nishigonaka, Myodaiji-cho, Okazaki, Aichi
444-8585, Japan
FEATURES
source Location/Qualifiers
1..3349
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="X"
/map="19E5-7"
916..2952
/note="TAK1 homolog"
/codon_start=1
/product="TGF-beta activated-kinase 1 homolog"
/protein_id="AAF06815.1"
/db_xref="GI:6289099"
/translation="MATASLDALQAAVVDSEITLRKVGHSYGVVCKAWRDKLVA
VKPEFSAEKOIDEKVKQLSRVPIALHGISYQOATYILMEABEGSLHNFJH
GKVPAYSLAHAMMAROCAGLAYLHAMPTKLIHSDVAPLILNKKRNLKIDF
GTVDKSTMTNNRNSGAAMWAPEVFGSKTEKDDISMAIIVMEVLSGRPKGIDN
AYTIQMTIKGERPPLLTTCPKRIEDLMTACWKTVPEDRSMQYIVMHEIVKDTYG
ADKALETFVNOQIVKESDGTVAQPDSSSQGEUSPSTQLTPTTANAVNAIA
ISKTTSMETNTSSTSDIPTNSGQLDNNPLFYVWVNTKMDALPEBESNESNDSE
LTSSEATFORLETRNGWILMACPKMQLTLDVANGFDLSPSSSSSTNAKSDGE
RLVTDFKPMVMTDLSNNNGIHAHNGLLSHANGQARDELEOGEHCOETLVNSID
VVDVDEDENGTEOSIASLILDELOPEKPIPNBASQLIYRHRHAKYKSLVDTNL
YYADPFKDKLIVQMDRTREBQKOLIKMKDKBQSLVNNLOQVASYRPLAGHHHQ
PHPHPHQLOHPSHPMHFLQDEGCGLLPSGVCGSGSESVEBGWVIPPHNHA"
BASE COUNT 934 a 893 c 853 g 669 t
ORIGIN
Alignment Scores:

Pred. No.: 5,51e-61 Length: 3349
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 3 Gaps: 3
US-09-830-144-2_COPY_76_303 (1-228) x AF199466 (1-3349)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsn11LeuValLeuTrpGly----- 19
Db 1093 GAGGTGAGACGATGTTGTGCGCCGTGAAGACCCGAACATCATCGCTTCCACGGGATATCC 1152
QY 20 AlaCysLeuAsnProValCysLeuValMetGluTrpAlaGluGlyGlySerLeuTrpAsn 39
Db 1153 TCGTACAGACAGCCACCTCATGATATATGATGATGATGCGGAGGTGATCGGCACMAC 1212
QY 40 ValLeuHisGlyAlaGluProLeuProTrpTrpThrAlaAlaHisAlaMetSerTrpCys 59
Db 1213 TTCCTTACGCGC--AAGGTGAAGCCGGCATATCTTCTGCGCCACCGCATGAGCTGGCG 1269
QY 60 LeuGlnCysSerGlnGlyValAlaTrpLeuHisSerMetGlnProValAlaLeuLeuHis 79
Db 1270 CGCCATGTGACAGAGGTCTGCGCATTTTGCATGCGCATGACCCGAAACCATTAATATCAT 1329
QY 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1330 CGCGACGTGAAGCGCGTGAACCTGCTTGTGACCAACAGGAGCGCAATCTGAAGATATGC 1389
QY 100 AspPheGlyThrAlaCysAspI11eGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db 1390 GACTTCGCGACGGTGGGGGACAGTGCACCATGATACCAACATGCGGTGTCGCGCT 1449
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTrpSerGluLysCysAspValPheSer 139
Db 1450 TGATGCGCGCCGACAGGTCTTGCAGAGCTCCAACTATACGAGAAAGTACATTTTCACG 1509
QY 140 TrpGlyLeuLeuLeuTrpGluValI11eThrArgArgLysProPheAspGluLeuGly 159
Db 1510 TGGCGCATTTGTTCTATGCGAGGTCTGTCGACAGAACACCCCTTTAAAGCATGCACAT 1569
QY 160 ProAlaPheArgI11eMetTrpAlaValHisAsnGlyThrArgProProLeuLeuLysAsn 179
Db 1570 --GCTACACCATTCAGTGAAGATCTACAGAGTGAAGCGCCGCGTGTGACCACT 1626
QY 180 LeuProLysPro11LeuSerLeuMetThrArgCysTrpSerLeuAspProSerGlnArg 199
Db 1627 TGCCCCAAGCGCATTCAGACACTGATGACCGCTGTGAAAAAGTGTCCAGAGATGCC 1686
QY 200 ProSerMetGluGluLeuValLysI11eMetTrpHisLeuMetArgTrpPheProGluAla 219
Db 1687 CCGTGTGATGACATATGATGCGGCTTATGACAGAGATGTCGCAAGACATATACGGGGCG 1746
QY 220 AspGluProLeuGlnTrp 225
Db 1747 GACAAAGCCCTGGAATAC 1764
RESULT 15
AY051953 3386 bp mRNA linear INV 27-AUG-2001
LOCUS Drosophila melanogaster LD42274 full length cDNA.
DEFINITION AY051953
ACCESSION AY051953
VERSION AY051953.1 GI:15292216
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pelecyota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 3386)
AUTHORS Stapleton, M., Brokstein, P., Hong, L., Agrawani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Rafan, D., Friese, E., George, R.,
Gonzalez, M., Guarin, H., Li, P., Liao, G., Miranda, A., Mungall, C. J.,

Nunoo,J., Pacieb,J., Paragas,V., Park,S., Phouanenvong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.

TITLE Direct Submission
JOURNAL Submitted (10-AUG-2001) Berkeley Drosophila Genome Project, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

COMMENT Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to sequence clones from Drosophila Gene Collection 1 (Rubin et al., Science 2000). The sequence has been subjected to integrity checks for sequence accuracy, presence of a polyA tail and contiguity within 100 kb in the genome. Thus we believe the sequence to reflect accurately this particular cDNA clone. However, there are artifacts associated with the generation of cDNA clones that may have not been detected in our initial analyses such as internal priming, priming from contaminating genomic DNA, retained introns due to reverse transcription of unspliced precursor RNAs, and reverse transcriptase errors that result in single base changes. For further information about this sequence, including its location and relationship to other sequences, please visit our Web site (<http://fruitfly.berkeley.edu>) or send email to cdna@fruitfly.berkeley.edu.

FEATURES

source

1..3386
/organism="Drosophila melanogaster"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/map="19E1-19E1"
1..3386
/gene="CG18492"
/note="alignment with genomic scaffold AB003571"
/db_xref="FLYBASE:FBgn0042230"
935..2971
/gene="CG18492"
/note="Longest ORF"
/codon_start=1
/product="LD4274p"
/protein_id="AAK93377.1"
/db_xref="GI:15292217"
/db_xref="FLYBASE:FBgn0042230"
/translation="WATASLDALQAYVDFSEITREKVGHSYGWVCKAVMRDKLVA
VKEFPASARQDIKEKVKOLSRVHPNIIALHGISYQOATVLIPEAEGGSLNHLFLH
GVKPAYSLAMSWARQACGLAYLHAWTPKPLIHRDVKPLNLLTNKGNLKICDF
GTVADKSTMTNNRGAAMWAEVFEFGKYTEKCDIFSWAIVLWEVLSRKQPFPGIDN
AYTIQWKIYKGERPPLITTCPKRIEDLMTACWKTVPEDRPSNQYIVGVNHEIVKDYTG
ADKALEYTFVNOQIVTKESDGTVAAPDSLSQEGELSPSSQTLPPTTAANANVIA
ISKTTTSSMTNTSSDITPTNSGQDNNPLFYVMTNRWDALPEESNESRNDSPN
LTSSAPATQRLRTIRNGMTLACKPMBOI.TLDVEANGFDLSPSESSSSSTNAKSDGRE
RLTVDTKVMWMTDLSNNNGIHAHNSGLLSHANGQARDEELQEQSHQEIYNSLD
VVDVDEDNTEQSLAIBLQEPPIFENDAESQLIYDRHRHMAKEYLSYVDVTLN
YYAODFKDLIVOMDTEREQKLELRMKQKGLQSLYNNLQQOYASRQKLAAGHPQ
PHPHPHNQLQHPHSHPPMHFLQDEGCGLLPGSVCGSGSEVBEGWVTPPHHNA"

gene

CDS

BASE COUNT 956 a 898 c 857 g 675 t
ORIGIN

Alignment Scores:

Pred. No.: 5.58e-61 Length: 3386
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 3 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x AY051953 (1-3386)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGly----- 19
Db 1112 GAGGTGAAGCAGTTGTCGCGCGTGAAGACCCGACCATCATCTGCTCGACGGGATATCC 1171
Qy 20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39

Db 1172 TCGTACCAGCAGGCCACCTACCTGATAATGAGTTCGCCGAAGGTGATCGCTGCACAAC 1231
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCys 59
Db 1332 TTCCTTCACGCG---AAGGTGAAGCCGCGCATATCTCTGCGCCACGCCCATGACGTGGCG 1288
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 1389 CGCCAATGTGCAGAGGCTCTGGCATAATTCATGTCATGACGCGCAAAACCACTAATACAT 1348
Qy 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1349 CGCACGTGAAGCCGCTGAACCTCTTGACCAACAAAGGCGCAATCTTGAAGATATGC 1408
Qy 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db 1409 GACTTCGGCAGCTGGCGGCAAGTCAGCATGATGACCAACCAATCGCGGAGTCGCCCT 1468
Qy 120 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSer 139
Db 1469 TGGATGGCGCCGAGGCTTCGAAGGCTCCAAGTATACGGAGAAGTGTGACATTTTATGC 1528
Qy 140 TrpGlyIlelleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 159
Db 1529 TGGGCCATTTGTTCTATGGAGGTTCTGTCCAGGAAGCAGCCCTTTAAAGGCATCGACAAT 1588
Qy 160 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 179
Db 1589 ---GCCTACACCATCCAGTGAAGATCTCAAGGTTGAACGCCCGCGCTGCTGACCACT 1645
Qy 180 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 199
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATGACCGCTGCTGGAACCGGTGCCGAGGATCGC 1705
Qy 200 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAla 219
Db 1706 CCGTCGATGCAGTACATAGTGGCGGTTATGACGAGATCGTCAAGGACTATACGGGGCG 1765
Qy 220 AspGluProLeuGlnTyr 225
Db 1766 GACAAGGCCCTTGAATAC 1783

Search completed: December 10, 2002, 04:57:50
Job time : 3132 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 10, 2002, 03:57:45 ; Search time 319 Seconds
(without alignments)

1609.579 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRYNHPNIVKLYGA.....MTHLMRYFFGADEPLQYPCQ 228

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US098301044/runat_04122002_141754_4332/app_query.fasta_1.391
-DB=N_geneseq_101002 -Qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US098301044 @CNG 1.1.79 @runat_04122002_141754_4332 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=10 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1252	100.0	1704	20	AAx99698	Human TGF-beta act
2	1252	100.0	1788	20	AAx56285	Human TAK1-6xHis e
3	1252	100.0	1959	18	AAT85095	Human transforming
4	1252	100.0	2443	18	AAT85094	Mouse transforming
5	1252	100.0	2656	20	AAx56279	Human TAK1 encodin
6	1252	100.0	2656	21	AAx39105	Human TAK-1 nucleo
7	1252	100.0	2769	24	ABL88437	Pain regulated cDN
8	1252	100.0	2785	20	AAx99696	Human TGF-beta act
9	1252	100.0	2866	20	AAx99697	Human TGF-beta act
10	700	55.9	3367	23	ABL02489	Drosophila melanog
11	599	47.8	10997	23	ABL02488	Drosophila melanog
12	414	33.1	759	23	ABL08337	Drosophila melanog
13	392.5	31.3	3454	24	ABL70018	Pancreas cancer re
14	374	29.9	1224	21	AAC43254	Arabidopsis thalia
15	373	29.8	1631	21	AAC39537	Arabidopsis thalia
16	371	29.6	1063	22	AAH34976	Human colon cancer
17	371	29.6	1365	21	AAZ99727	cDNA encoding huma
18	371	29.6	1706	21	AAx75674	DNA encoding a hum
19	371	29.6	2120	21	AAZ99726	cDNA encoding huma
20	371	29.6	2191	22	AAH99263	Human protein enco
21	371	29.6	2194	21	AAZ99734	Cardiovascular sys
22	371	29.6	2220	21	AAZ93783	Human survival reg
23	371	29.6	2254	21	AAZ99735	Cardiovascular sys
24	371	29.6	2403	22	AAx44701	Novel protein kina
25	371	29.6	2622	22	AAx75336	Human TGF-beta rec
26	371	29.6	3967	22	AAH73366	Human cervical can
27	370	29.6	2069	21	AAZ99737	Cardiovascular sys
28	370	29.6	2272	21	AAZ99736	Cardiovascular sys
29	363.5	29.0	3072	23	ABL04365	Drosophila melanog
30	361.5	28.9	3141	22	AAD18824	Human kinase (PKIN
31	361.5	28.9	3538	24	AAD34309	Human PKIN-12 CDNA
32	355.5	28.4	3066	24	ABQ86165	Novel human gene.
33	355	28.4	2157	22	AAH46913	cDNA encoding huma
34	355	28.4	3111	24	ABN86357	Novel human protei
35	355	28.4	3518	24	ABN86358	Novel human protei
36	353.5	28.2	3365	24	ABK84203	Human cDNA differe
37	353.5	28.2	3389	16	AAT01031	Human leucine zipp
38	353.5	28.2	3389	18	AAT89349	Human leucine-zipp
39	352.5	28.2	3427	24	AB199250	Mouse ischaemic co
40	351.5	28.1	3109	23	ABL29755	Drosophila melanog
41	350	28.0	2283	21	AAC48526	Arabidopsis thalia
42	348	27.8	1591	21	AAC40839	Arabidopsis thalia
43	347	27.7	1659	21	AAC50138	Arabidopsis thalia
44	347	27.7	1661	21	AAC41921	Arabidopsis thalia
45	345.5	27.6	3558	24	ABK83874	Human cDNA differe

ALIGNMENTS

RESULT 1

AAx99698

ID AAx99698 standard; cDNA to mRNA; 1704 BP.

XX

AC AAx99698;

XX

DT 18-OCT-1999 (first entry)

XX

DE Human TGF-beta activated kinase (TAK) ic encoding nucleotide sequence.

XX

Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
intractable disease; atrophic dermatitis; psoriasis; viral infection;
endotoxin shock; septicemia; human; hTAK1c; ss.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT	CDS	1..1704
FT	1	/*tag= a
FT	product	= "hTAK1c"
PN	WO9940202-A1.	
XX		
PD	12-AUG-1999.	
XX		
PF	02-FEB-1999;	99WO-JP00422.
XX		
PR	30-OCT-1998;	98UP-0309316.
PR	06-FEB-1998;	98UP-0026003.
XX		
PA	(TANA) TANABE SEIYAKU CO.	
XX		
PI	Haasegawa K, Kageyama N, Sakurai H, Sugita T;	
XX		
DR	WPI; 1999-494298/41.	
XX		
DR	P-PSDB; AAY28998.	
PT	Nuclear factor kappa B activation inhibitors, useful as preventives	
XX	for, e.g. autoimmune diseases	
PS	Examples; Page 43-46; 49pp; Japanese.	
XX		
CC	The invention provides a method for identifying or screening a nuclear	
CC	factor kappa B (NF-kB) activation inhibitor by examining the effect of a	
CC	factor kappa B (NF-kB) activation inhibitor by examining the effect of a	
CC	test substance on modulating the function(s) of TGF-beta activated kinase	
CC	1 (TAK1). The Nk-B activation inhibitors targeting on TAK1 can be used to	
CC	treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis)	
CC	intractable diseases with inflammation (such as atrophic dermatitis and	
CC	psoriasis), viral infection, endotoxin shock, septicemia and others. The	
CC	present sequence represents the nucleotide sequence of human TAK1c	
CC	(hTAK1c) protein.	
XX		
XX	Sequence 1704 BP; 511 A; 381 C; 401 G; 411 T; 0 other;	

Alignment Scores:	
Pred. No.:	4,71e-133
Score:	1252.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	20
US-09-830-144-2_COPY_76_303 (1-228) x AAX96698 (1-1704)	
	Length: 1704
	Matches: 228
	Conservative: 0
	Mismatches: 0
	Indels: 0
	Gaps: 0

Oy	1	VaGluLeuAArgGlnLeuSerAtrValAsnHisproAsnLeuVallySLeuYrGlyAla	20
Db	226	GTAAAGCTTCGGCAGTTATCCCGGTGAACATCTTAATTTGAAAGCTTTATGGAGCC	285
Oy	21	CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuYrAsnVal	40
Db	286	TGCTTGAAATCCAGATGCTGCTTGATGGAATATCTTAAGGGGGCTTTATATTAATGTG	345
Oy	41	LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu	60
Db	346	CTGCATGATGCTGAACCATTCCTCATATTATTAATCTCTGCCCCAGCCAAATGATGGTGTTHA	405
Oy	61	GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg	80
Db	406	CAGTGTCTCCAAAGGAGTGGCTTAATCTTCACAGCATGCAACCCAAAGGCGTAATTTCACAG	465
Oy	81	AspLeuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp	100
Db	466	GACCTGAAACCAACCAATTTACTCTGCTGTGCAGGGGGAGCAAGTTCTTAAATAATTTGGAT	525
Oy	101	PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyr	120
Db	526	TTTGTTAACAGCTGTGACATTCACAGCACACATGACCAATTAACAAAGGGAGATGCTGCTTGG	585
Oy	121	MetAlaProGluValPheGlyGlySerAsnTyrSerGluIlyCysAspValPheSerTyr	140

Db	586	ATGGCACTGAAAGTTTGTGAAGTGAATTAATCAAGTGAATAATGTACGCTTTCAGCTGG	645
Qy	141	GLYIleIleuLeuTPGjLuValIleThArGrghYsProPheAspGjuliLegIYjPro	160
Db	646	GGTATTATTCTTTGGGAAGCATTAACGGTGGAAACCTTTGATGAGATTGGTGCCCA	705
Qy	161	AlAPheArgIleWetTPpAlaValHisnGlyThiArpProPoleuIleLysnLeu	180
Db	706	GCTTCCGAATCAAGTGGCGCTGTCAATATGGTACTGCACACACACATGATAAAAATTTA	765
Qy	181	ProLysProIleGjLeuSerLeuMetThiArGrCysTrpSerLysAspProSerGjlnArGrPro	200
Db	766	CTTAAGCCCATTTGAGAGCGCTATGACTGCTGTTGGTTAAGATCCTTCCAGAGGCCCT	825
Qy	201	SerMetGjLuIleValLysIleWetThiHisLeuMetArgIlyrPheProGjlnAlaAsp	220
Db	826	TCAATTGGAGGAAATTGTGTAATAATGATCATCTTGATGACCGTACTTCCAGAGACAGAT	885
Qy	221	GLuProLeuGjlnIlyrProCysGjln	228
Db	886	GAGCATTTACAGTATCCTTGTACG	909

XX	AAK56285	standard; DNA; 1788 BP.
ID	AAK56285	
AC	AAK56285;	
XX		
DT	21-JUL-1999	(first entry)
XX		
DE	Human TAK1-6xHis encoding DNA.	
XX		
KW	Human; TAB1; TAK1; screening; inhibition; TGF-beta;	
KW	transforming growth factor beta; ss.	
XX		
OS	Homo sapiens.	
XX	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	7..1779	
FT	CDS	/*tag= a
XX		
PN	WO9921010-A1.	
XX		
PD	29-APR-1999.	
XX		
PF	22-OCT-1998;	98WO-JP04796.
XX		
PR	22-OCT-1997;	97JP-0290188.
XX		
PA	(CHUS) CHUGAI SEIYAKU KK.	
XX		
PI	Ohomo T, Ono K, Tsuchiya M;	
XX		
DR	WPI; 1999-312645/26.	
DR	P-PSDB; AAY05547.	
XX		
PT	Screening for TGF- beta inhibitory substances, which are useful as	
PT	drugs for treatment of diseases relating to its disorder	
XX		
XX	Example 1, Page 167-171; 195pp; Japanese.	
XX		
XX	A method has been developed for screening for substances which inhibit the binding of TAK1 polypeptide to TAB1 polypeptide. The method comprises: (a) contacting the polypeptide in the presence of a sample; and (b) detecting the amount of bound polypeptide, in which the sample can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming growth factor (TGF)-beta inhibitory substances can be used in drugs for indications e.g. as TGF-beta signal transmission inhibitors or activators, or extracellular matrix protein production enhancement inhibitors or activators, or cell proliferation prevention inhibitors or activators, or monocyte migration inhibitors or activators, or physiological activity induction inhibitors or activators, or	
CC		

CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes TAK1-6xHis from an example of
CC the present invention.

XX
SQ Sequence 1788 BP; 551 A; 394 C; 413 G; 430 T; 0 other;

Alignment Scores:
Pred. No.: 5,04e-133 Length: 1788
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAX56285 (1-1788)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 232 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 291
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 292 TGCCTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 351
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 352 CTGCATGTGTCTGAACCATTTGCATATTATCTGCTGCCACGCAATGAGTTGGTGTITA 411
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 412 CAGTGTCTCCCAAGGAGTGGCTTATCTTCACAGCATGCACCAACCAAGCGCTAATTCACAGG 471
QY 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 472 GACCTGAACACCAACTTACTGCTGGTTCAGGGGGGACAGTTCTAAAAATTTGTGAT 531
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
DB 532 TTTGGTACAGCTGTGACATTCAGACACATGACCATTAACCAAGGGAGTCTGCTTGG 591
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
DB 592 ATGGCACCTGAAGCTTTTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 651
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysPheAspGluIleGlyPro 160
DB 652 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAAACCCCTTTGATGAGATTGGTGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProLeuIleLysAsnLeu 180
DB 712 GCCTTCGAATCATGTGGCTGTTCTATATGTTGTTCTGCACCATCTGATAAATAATTTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 772 CCTAAGCCCATTCAGAGCCTGATGACTCGTTGTTGTCTTAAAGATCCTTCCAGCGCCCT 831
QY 201 SerMetGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
DB 832 TCAATGAGGAAATGTGAAAAATATGACTCACTTGTATGGGTACTTTTCCAGGAGCAGAT 891
QY 221 GluProLeuGlnTyrProCysGln 228
DB 892 GAGCCATTACAGTATCCTTGTCTCAG 915

RESULT 3

ID AAT85095 standard; cDNA; 1959 BP.

XX AAT85095;

AC AAT85095;

XX 19-NOV-1997 (first entry)

XX Human transforming growth factor-beta activated kinase TAK-1 cDNA.
DE
XX TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
KW protein kinase; ss.
XX Homo sapiens.

XX Key Location/Qualifiers
FH CDS 183..1922
FT /*tag= a
FT /product= TAK-1
XX
XX JP09163990-A.

XX
XX 24-JUN-1997.
XX
XX 27-SEP-1996; 96JP-0256747.
XX
XX 24-JUL-1996; 96US-0685625.
PR 29-SEP-1995; 95JP-0253549.
XX
XX (CHUS) CHUGAI PHARM CO LTD.
PA (UENO//) UENO N.

XX
XX WPI; 1997-380171/35.
DR P-PSDB; AAW27093.
XX

PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
PT - useful for studying the TGF-beta signal transduction system
XX
XX Claim 9; Page 13-15; 20pp; Japanese.
XX
XX The present sequence encodes human transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal
CC transduction system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
XX
SQ Sequence 1959 BP; 571 A; 447 C; 481 G; 460 T; 0 other;

Alignment Scores:
Pred. No.: 5,74e-133 Length: 1959
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAT85095 (1-1959)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAAGCTTTATGGAGCC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
DB 468 TGCCTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 528 CTGCATGTGTCTGAACCATTTGCATATTATCTGCTGCCACGCAATGAGTTGGTGTITA 587
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 588 CAGTGTCTCCCAAGGAGTGGCTTATCTTCACAGCATGCACCAACCAAGCGCTAATTCACAGG 647
QY 81 AspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 648 GACCTGAACCAACTTACTGCTGTTGACGGGGGACAGTTCTAAAAAATTTGTGAT 707

```

OY 101 PheGlyThrAlaCysasp11legInThrHisMetThrAsnAsnLysglYSerAlaIaIaTrp 120
DB 708 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATACAGGGAGGTGCTGCTGG 767
OY 121 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 140
DB 768 ATGGCACCTGAAGTTTGTGAAGTAGTAATTAACAGTGAAGAAATGTGACGCTTCAGCTGG 827
OY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
DB 828 GGTATTATTCTTTGGAGAGATACGCGTCGAAACCTTTGATGAGATTGTGTGCCCA 887
OY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 888 GCTTTCGAAATCATGTGGGCTGTCTAATGTGACTCGACCCACCATGATTAATAATTTA 947
OY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 948 CTTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGGTCTTAAGATCTTCCAGGCGCCT 1007
OY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyRheProGlyAlaAsp 220
DB 1008 TCAATGGAGAAATGTGAAAATTAATGACTCACTTGATCGGTACTTCCAGAGCAGAT 1067
OY 221 GluProLeuGlnTyRProCysGln 228
DB 1068 GAGCCATTACAGTATCTTGTCTG 1091

RESULT 4
AAT85094
ID AAT85094 standard; cDNA; 2443 BP.
XX
AC AAT85094;
XX
DT 19-NOV-1997 (first entry)
XX
DE Mouse transforming growth factor-beta activated kinase TAK-1 cDNA.
XX
KM TGF-beta; signal transduction; TGF-beta activated kinase;
KW MAPK kinase activator; AMK-1; bone morphogenetic protein; BMP;
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT CDS 157..1896
FT FT /*tag= a
FT FT /product= TAK-1
XX
PN JP09163990-A.
XX
PD 24-JUN-1997.
XX
PF 27-SEP-1996; 96JP-0256747.
XX
PR 24-JUL-1996; 96US-0695625.
PR 29-SEP-1995; 95JP-0235349.
XX
PA (CHUS) CHUGAI PHARM CO LTD.
PA (UENO) UENO N.
XX
DR MPI; 1997-380171/35.
DR P-PSDB; AAM27092.
XX
PT DNA encoding transforming growth factor-beta-activated kinase, TAK-1
XX - useful for studying the TGF-beta signal transduction system
XX
PS Claim 2; Page 10-12; 20pp; Japanese.
XX
CC The present sequence encodes mouse transforming growth factor-beta
CC (TGF-beta) activated kinase, TAK-1. The DNA is used to produce the
CC TAK-1 protein which is involved in the TGF-beta family signal

```

```

CC transmission system. TAK-1, also known as activator of MAPK Kinase
CC (AMK-1), is an enzyme which is activated by TGF-beta and bone
CC morphogenetic protein (BMP) and activates MAPK kinase by
CC phosphorylation.
XX
SQ Sequence 2443 BP; 669 A; 567 C; 647 G; 559 T; 1 other;
Alignment Scores:
Pred. No.: 7 86e-133 Length: 2443
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 18 Gaps: 0
US-09-830-144-2_COPY_76_303 (1-228) x AAT85094 (1-2443)
OY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyRGIYAla 20
DB 382 GTGAGCTCCGGGAGTGTGCGGTGTGAACCATCTTAACATTTGTCAAGTTGTACGAGCC 441
OY 21 CysLeuAsnProValCysLeuValMetGluTyRAlaGluGlySerLeuTyRAsnVal 40
DB 442 TGCCGTAATCCAGATGTCTTGTGATGGAATATGACAGAGGGGCGCTCATTTGTAATGTG 501
OY 41 LeuHisGlyValGluProLeuProTyRTrpThrAlaAlaHisIleMetSerTrpCysLeu 60
DB 502 CTGCATGTGTCTGAACCATTTGCTCTTACTACATGCTGCTCAAGCCATGAGCTGTGTTTA 561
OY 61 GlnCysSerGlnGlyValAlaTyRLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 562 CAGTGTTCCCAAGAGAGTGGCTTACCTGCACAGCATCGACGCCCAACCGCTGATTCACAG 621
OY 81 AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 622 GACCTCAAGCCTCCAACTTGCTGTGTTGCAGAGGAGCAGTCTTAATAATCTGCAT 681
OY 101 PheGlyThrAlaCysasp11legInThrHisMetThrAsnAsnLysglYSerAlaIaIaTrp 120
DB 682 TTTGGTACAGCCTGTGACATTCACCAACACATGACCAATTAAGGAGTGTCTGCTGG 741
OY 121 MetAlaProGluValPheGluGlySerAsnTySerGluLysCysAspValPheSerTrp 140
DB 742 ATGGCGCCTGAAGTGTGAGAGTAGCAATTAACAGTGAAGTGTATGCTTCAGCTGG 801
OY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
DB 802 GGTATTATCTCTGGGAGATGATACACGCCGGAACCTTCGATGAGATCGGTGCCCA 861
OY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
DB 862 GCTTTCGAAATCATGTGGGCTGTCTAATTAAGCAGCTGACACCATGATTAATAATTTTA 921
OY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
DB 922 CTTAAGCCCATTTGAGAGCTTGAAGTACACGCTGTGTTAAGACCACTCAGCGCCT 981
OY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyRheProGlyAlaAsp 220
DB 982 TCAATGAGAGAAATGTGAAAATTAATGACTCACTTATCGGTATCCAGAGAGGAT 1041
OY 221 GluProLeuGlnTyRProCysGln 228
DB 1042 GAGCCATTACAGTATCTTGTCTG 1065

RESULT 5
AAX56279
ID AAX56279 standard; DNA; 2656 BP.
XX
AC AAX56279;
XX
DT 21-JUL-1999 (first entry)
XX

```

DE Human TAK1 encoding DNA.

XX Human; TAB1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FH 183...1922
FT CDS /*tag= a

XX WO9921010-A1.

XX 29-APR-1999.

XX 22-OCT-1998; 98WO-JP04796.

XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.

XX Ohtomo T, Ono K, Tsuchiya M;

XX WPI; 1999-312645/26.

DR P-PSDB; AAY09542.

XX Screening for TGF- beta inhibitory substances, which are useful as
PT drugs for treatment of diseases relating to its disorder

PS Example 1; Page 150-154; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAK1.

XX Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;

SQ Alignment Scores:

Pred. No.: 8.85e-133 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAX56279 (1-2656)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGAACCATCTTAATATTGTAAGCTTTATGGAGCC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyCysLeuTyrAsnVal 40
Db 468 TGCATGAATCCAGTGTCTCTGTGTGAATATGCTGAAGGGGGCTCTTATATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGCATGTGCTGAACCATTTGCCATATTATCTGTGCGCCACCAATGAGTTGGTGTATA 587
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80

Db 588 CAGTGTTCCTCCCAAGGAGTGGCTTATCTTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 647
QY 81 AspLeuIysProProAsnLeuLeuValAlaGlyClyThrValLeuIysIleCysAsp 100
Db 648 GACCTGAACCAACCAACTTACTGCTGTGTTGCAGGGGGACAGTTCTAAAAATTTGTGAT 707
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyr 120
Db 708 TTTGGTACAGCGCTGTGACATTCAGACACACATGACCAATAACAAGGGGAGTGTGCTTGG 767
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerCLeuIysCysAspValPheSerTrp 140
Db 768 ATGCACCTGAAGTTTGAAGGTAGTAATTACAGTGAATAATGTGACGCTCTTCAGCTGG 827
QY 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 828 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAACCCCTTTGATGAGATTGGTGGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db 888 GCTTTCGAATCATGTGGCTGTTCTAATAATGTTACTCGACCACCATGATAAAAAATTTA 947
QY 181 ProIysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 948 CCTAAGCCCATTTGAGAGCCTGATGACTCGTTGTGTGTTAAAGATCCTCCACGCGCCT 1007
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db 1008 TCAATGGAGGAATTGTGAATAATAATGACTCATTGATGGGTACTTTCCAGGAGCAGAT 1067
QY 221 GluProLeuGlnTyrProCysGln 228
Db 1068 GAGCCATTACAGTATCTTGTGCAG 1091
RESULT 6
AAA39105
ID AAA39105 standard; DNA; 2656 BP.
XX AAA39105;
AC AAA39105;
XX 04-SEP-2000 (first entry)
DT Human TAK-1 nucleotide sequence SEQ ID NO:1.
XX Human TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
DE screening; signal transduction; inhibition; inflammatory cytokine;
XX IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression; ds.
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT CDS 183...1922
FT /*tag= a
FT /product= "TAK-1"
XX WO200023610-A1.
XX 27-APR-2000.
XX 21-OCT-1999; 99WO-JP05817.
XX 21-OCT-1998; 98JP-0299962.
XX (CHUS) CHUGAI SEIYAKU KK.
XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
PI WPI; 2000-339707/29.
DR P-PSDB; AAY91000.
XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as

PT antiinflammatory agents -
XX
PS Example 1, Page 73-80; 100pp; Japanese.
XX
CC The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug
CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as Interleukin-1 (IL-1),
CC IL-10, tumor necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence encodes human TAK-1, which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 2656 BP; 778 A; 557 C; 620 G; 701 T; 0 other;

Alignment Scores:
Pred. No.: 8,85e-133 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AAA39105 (1-2656)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIlyLeuTyrgIyala 20
DB 408 GTAGAGCTTCGGCAGATTATCCCGTGGAACCATCTTAATATTGTAAGCTTTATGAGACC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40
DB 468 TGGTTGAATCCAGTGTGCTTTGTGATGAAATATGCTGMAAGGGGGCTCTTTATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyTyThraAlaAlaHisAlaMetSerTyPCysLeu 60
DB 528 CTGCATGTGTGTCGAACCATTTGCCATATTATCTGCTGCCACGATGATGGTGGTGTTA 587
QY 61 GlnCysSerGlnGlyValAlaTyrlleuHisSerMetGlnProIyAlaLeuIleHisArg 80
DB 588 CAGTGTTCCTCCAGAGAGTGTCTTATCTTCAACCATGCAACCCAAAGCCGTATATTCACAG 647
QY 81 AspLeuIyProProAsnLeuLeuValAlaGlyGlyThraValIleuIyIleCysAsp 100
DB 648 GACCTGMAACCAACCAACTTACTGCTGTGTCAGGGGGGACAGTTCTMAAAATTTGTGAT 707
QY 101 PheGlyThraIaCysAspIleGlnThraHisMetThraAsnAsnGlySerAlaAlaTrp 120
DB 708 TTTGGTACAGCGCTGTGACATTCAGACACACATGACCAATACAAAGGGAGTGTCTGTGG 767
QY 121 MetAlaProGluValPheGluGlySerAsnTyIySerGluIyIyCysAspValPheSerTrp 140
DB 768 ATGGCACCTGGAAGTTTAAAGTAGTAATTAACGTGAAAATGTGAGCTTCCAGCTGG 827
QY 141 GlyIleIleLeuTrpGluValIleThraArgArgIyProPheAspGluIleGlyIyPro 160
DB 828 GGTATTATTCTTTGGGAAGTGAATACGCGTCGGAACCTTTTGAAGAATGGTGGGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThraArgProProLeuIleIyAsnLeu 180
DB 888 GCTTCCGATCATGTGGGCTGTTCATAATGTGTACTCGACCCACCTGATTAATAATTTTA 947
QY 181 ProIySProlleGlySerleuMetThraArgCysTrpSerIyAspProSerGlnArgPro 200
DB 948 CCTAAGCCCATTAAGAGCCTGTATGACTCGTTGTGTCTTAAGATCTTCCACAGGCCCT 1007
QY 201 SerMetGluGluIleValIyIleMetThraHisleuMetArgTyIyPheProGlyAlaAsp 220

DB 1008 TCATGAGAGAAATTGCAAAATATATGACTTGAATGCGGTACTTTCAGAGAGAGAT 1067
QY 221 GluProLeuGlnTyProCysGln 228
DB 1068 GAGCCATTACAGTATCTTGTGAG 1091
RESULT 7
ABL88437
ID ABL88437 standard; cDNA; 2769 BP.
XX
AC ABL88437;
XX
DT 16-MAY-2002 (first entry)
XX
DE Pain regulated cDNA sequence 80.
XX
KM Pain; analgesic; gene therapy; neurological disorder;
KW neurodegenerative disease; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212338-A2.
XX
PD 14-FEB-2002.
XX
PF 03-AUG-2001; 2001WO-EP09011.
XX
PR 03-AUG-2000; 2000DE-1037759.
XX
PA (CHEF) GRUENENTHAL GMBH.
XX
PI Gillen C, Wetzel S, Wendt S, Welte E, Schaefer MK;
XX
DR WPI; 2002-257469/30.
XX
P-PSDB; ABB85033.
XX
PT Identifying pain-regulating compounds, useful for treating chronic pain
PT and for diagnosis, by measuring binding of compounds to specific
PT peptides and proteins -
XX
PS Claim 1, Fig 44; 213pp; German.
XX
CC The invention relates to identifying pain-regulating substances (A)
CC comprises (1) incubating a test substance with a cell (or preparation
CC from it) that has synthesized a peptide or protein (B) and (ii) measuring
CC either binding of the test substance to (B) or some functional parameter
CC that is altered by this binding. The method is useful for identifying
CC pain-regulating substances (A) with analgesic activity (A) along with
CC nucleic acid (ABL88411-ABL88441) that encode proteins (B).
CC ABB85006-ABB85037) that interact with (A); (B); vectors containing the
CC nucleic acid; antibodies against (B); cells that express (B) and agents
CC that bind to (B), are all useful for treating pain, particularly chronic
CC pain, including use in gene therapy. The same materials can also be used
CC for diagnosis e.g. of neurological and neurodegenerative diseases. The
CC present sequence is that of a polynucleotide of the invention.
XX
SQ Sequence 2769 BP; 811 A; 565 C; 640 G; 753 T; 0 other;

Alignment Scores:
Pred. No.: 9,39e-133 Length: 2769
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 24 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x ABL88437 (1-2769)
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIlyLeuTyrgIyala 20
DB 388 GTAGAGCTTCGGCAGATTATCCCGTGGAACCATCTTAATATTGTAAGCTTTATGAGACC 447
QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40

```

Db      448 TGCTTGAATCAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATATATGTG 507
Qy      41  LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGAACCATTTGCCATATTATCTACTGCTCCACCAATGAGTTGGTGTTTA 567
Qy      61  GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      568 CAGTGTTCCTCCCAAGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 627
Qy      81  AspleuLysProProLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      628 GACCTGAAACCCAACTTACTGCTGGTTGCGAGGGGGACAGTTCTTAAAAATTTGTGTAT 687
Qy      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db      588 TTGGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAAGGGGAGTGTGCTTGG 747
Qy      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db      748 ATGCGACCTGAAGTTTGTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy      141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyPro 160
Db      808 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAAACCCCTTTGATGAGATTGGTGCCCA 867
Qy      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
Db      868 GCITTTCCGAATCATGTGGGCTGTTCAATAGGTACTCGACCACCATGTATAAAAAATTTA 927
Qy      181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db      928 CCTAAGCCCATTTGAGACCTGATGACTCGTTGTTGTCTTAAAGATCCTTCCAGCGCCCT 987
Qy      201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
Db      988 TCAATGGAGGAATTTGAAAAATAATGACTCACTTGTATGCGGTACTTTCCAGGAGCAGAT 1047
Qy      221 GluProLeuGlnTyrProCysGln 228
Db      1048 GAGCCATTACAGTATCCTCTTGTCTAG 1071

RESULT 8
AAAX99696
ID      AAAX99696 standard; cDNA to mRNA; 2785 BP.
XX
AC      AAAX99696;
XX
XX
DT      18-OCT-1999 (first entry)
XX
DE      Human TGF-beta activated kinase (TAK) 1a encoding nucleotide sequence.
XX
KW      Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
KW      TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
KW      intractable disease; atrophic dermatitis; psoriasis; viral infection;
KW      endotoxin shock; septicemia; human; hTAK1a; ss.
XX
OS      Homo sapiens.
XX
XX      Key      Location/Qualifiers
FH      163..1902
FT      CDS      /*tag= a
FT      /*product= "hTAK1a"
XX
XX      WO9904202-A1.
XX
XX      12-AUG-1999.
XX
XX      02-FEB-1999; 99WO-JP00422.
XX
XX      30-OCT-1998; 98JP-0309316.
XX      06-FEB-1998; 98JP-0026003.

```

```

XX      (TANA ) TANABE SEIYAKU CO.
XX
XX      Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX
XX      WPI; 1999-494298/41.
DR      P-PSDB; AAY28996.
XX
XX      Nuclear factor kappa B activation inhibitors, useful as preventives
XX      for, e.g. autoimmune diseases
XX
XX      Examples; Page 35-39; 49pp; Japanese.
XX
XX      The invention provides a method for identifying or screening a nuclear
XX      factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX      test substance on modulating the function(s) of TGF-beta activated kinase
XX      1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX      treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX      intractable diseases with inflammation (such as atrophic dermatitis and
XX      psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX      present sequence represents the nucleotide sequence of human TAK1a
XX      (hTAK1a) protein.
XX
XX      Sequence 2785 BP; 827 A; 565 C; 640 G; 753 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 9.47e-133 Length: 2785
XX      Score: 1252.00 Matches: 228
XX      Percent Similarity: 100.00% Conservatives: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 20 Gaps: 0
XX
XX      US-09-830-144-2_COPY_76_303 (1-228) x AAAX99696 (1-2785)

```

```

Qy      1  ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
Db      388 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTTAATATTGTAAGAGCTTTATGGAGCC 447
Qy      21  CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsnVal 40
Db      448 TGCITGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 507
Qy      41  LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
Db      508 CTGCATGCTGCTGAACCATTTCCCATATTATCTGCTGCCACGCAATCAGTTGTTGTTA 567
Qy      61  GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
Db      568 CAGTGTTCCTCCCAAGGAGTGGCTTATCTTCACAGCATGCAACCCAAAGCGCTAATTCACAGG 627
Qy      81  AspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
Db      628 GACCTGAAACCCAACTTACTGCTGGTTGCGAGGGGGACAGTTCTTAAAAATTTGTGTAT 687
Qy      101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db      688 TTTGGTACAGCTGTGACATTCAGACACATGACCAATAACAAGGGAGTGTGCTTGG 747
Qy      121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
Db      748 ATGCGACCTGAAGTTTGTGAAGGTAGTAATTACAGTGAAAAATGTGACGTCTTCAGCTGG 807
Qy      141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db      808 GGTATTATTCTTTGGGAAGTGATAACGCTCGGAAACCCCTTTGATGAGATTGGTGCCCA 867
Qy      161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuLysAsnLeu 180
Db      868 GCITTTCCGAATCATGTGGGCTGTTCAATAGGTACTCGACCACCATGTATAAAAAATTTA 927
Qy      181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200

```

```
Db 928 CCHAAACCCATTGAGAGCCTGATGACTCGTGTGTGCTTAAGATCCTTCCAGGCGCCT 987
Qy 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyRpheProGlyAlaasp 220
Db 988 TCATATGAGGAAATGTGAAAATAATGACTCACTTGATCGGTACTTCCAGAGCAGAT 1047
Qy 221 GluProLeuGlnTyRProCysGln 228
Db 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071

RESULT 9
AAx99697
ID AAx99697 strand: cDNA to mRNA; 2866 BP.
XX
XX AAx99697;
XX
XX 18-OCT-1999 (first entry)
XX
DE Human TGF-beta activated kinase (TAK) 1b encoding nucleotide sequence.
XX
XX Nuclear factor kappa B; NF-kB; inhibitor; TGF-beta activated kinase 1;
XX TAK1; autoimmune disease; chronic rheumatoid arthritis; inflammation;
XX intracetable disease; atrophic dermatitis; psoriasis; viral infection;
XX endotoxin shock; septicemia; human; hTAK1b; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT CDS 163..1983
XX FT /tag= a
XX FT /product= "hTAK1b"
XX
XX WO9940202-A1.
XX
XX 12-AUG-1999.
XX
XX 02-FEB-1999; 99WO-JP00422.
XX
XX 30-OCT-1998; 98JP-0309316.
XX PR 06-FEB-1998; 98JP-0026003.
XX
XX (TANA ) TANABE SEIYAKU CO.
XX
XX Hasegawa K, Kageyama N, Sakurai H, Sugita T;
XX
XX WPI; 1999-494298/41.
XX DR P-PSDB; AAy28997.
XX
XX Nuclear factor kappa B activation inhibitors, useful as preventives
XX PT for, e.g. autoimmune diseases
XX
XX Examples; Page 39-43; 49pp; Japanese.
XX
XX The invention provides a method for identifying or screening a nuclear
XX factor kappa B (NF-kB) activation inhibitor by examining the effect of a
XX test substance on modulating the function(s) of TGF-beta activated kinase
XX 1 (TAK1). The NFkB activation inhibitors targeting on TAK1 can be used to
XX treat or prevent autoimmune diseases (e.g. chronic rheumatoid arthritis),
XX intracetable diseases with inflammation (such as atrophic dermatitis and
XX psoriasis), viral infection, endotoxin shock, septicemia and others. The
XX present sequence represents the nucleotide sequence of human TAK1b
XX (hTAK1b) protein.
XX
XX Sequence 2866 BP; 848 A; 588 C; 656 G; 774 T; 0 other;

Alignment Scores:
Pred. No.: 9,86e-133 Length: 2866
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0
```

```
US-09-830-144-2_COPY_76_303 (1-228) x AAx99697 (1-2866)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyRGIyAla 20
Db 388 GTAGAGCTTGCGGAGATTATCCCGGTGAGAACCATCTTAATATGTAAGCTTTATGAGGCC 447
Qy 21 CysLeuAsnProValCysLeuValIleMetGluTyRAlaGlnGlySerLeuTyRAsnVal 40
Db 448 TGCCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTG 507
Qy 41 LeuHisGlyAlaGluProLeuProTyRThrAlaAlaHisAlaMetSerTyRProCysLeu 60
Db 508 CTGATGTGTCTGACCATTCATGTCATATATATACGCGCCACGCCATGATGTTGTCTTTA 567
Qy 61 GlnCysSerGlnGlyValAlaTyRLeuHisSerMetGlnProIlyAsnAlaLeuIleHisArg 80
Db 568 CAGTGTTCACAGAGAGTGGCTTATCTTCACAGCATCAACCCAAACGGCTAATTCACAG 627
Qy 81 AspleuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
Db 628 GACCTGAAACCCAACTTACTGTGTGTGACGGGGGACAGTTCTTAATAATTTGTGAT 687
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyR 120
Db 688 TTTGTATCAGCCTGTGACATTCAGACACATGACCAATTAACAGGGGAGTGTCTGTGG 747
Qy 121 MetAlaProGluValPheGlnIysSerAsnTyRSerGluIysCysAspValPheSerTyR 140
Db 748 ATGGCACCCTGAAGTTTGTGAAGGTAGTAATTAACAGTGAATAAGTGAACGCTTCACGCTGG 807
Qy 141 GlyIleIleLeuTyRArgValIleThrArgArgIysProPheAspGlnIleGlyGlyPro 160
Db 808 GGTATTTATCTTTGGGAAAGATPACCGCTGGGAAACCTTTGATATAGATGGTGGCCCA 867
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
Db 868 GCTTCCGAATCATGTGGCTGTTCATTAATGTACTCGACACCATGATTAATAAATTTA 927
Qy 181 ProIysProIleGlnSerLeuMetThrArgCysTrpSerIysAspProSerGlnArgPro 200
Db 928 CCHAAACCCATTGAGAGCCTGATGACTCGTGTGTGCTTAAGATCCTTCCAGCGCCCT 987
Qy 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyRpheProGlyAlaasp 220
Db 988 TCATATGAGGAAATGTGAAAATAATGACTCACTTGATCGGTACTTCCAGAGCAGAT 1047
Qy 221 GluProLeuGlnTyRProCysGln 228
Db 1048 GAGCCATTACAGTATCCTTGTCTCAG 1071

RESULT 10
ABL02489
ID ABL02489 strand: cDNA; 3367 BP.
XX
XX ABL02489;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SHQ ID NO 1949.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
```

```

XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58386.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1949; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 3367 BP; 938 A; 898 C; 856 G; 675 T; 0 other;

Alignment Scores:
Pred. No.: 1.77e-69 Length: 3367
Score: 700.00 Matches: 130
Percent Similarity: 73.01% Conservative: 35
Best Local Similarity: 57.52% Mismatches: 57
Query Match: 55.91% Indels: 4
DB: 23 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x ABL02489 (1-3367)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrgly----- 19
Db 1112 GAGTGAAGCAGTGTTCGCGGTGAACACCCGAAACATCATCGCTTGACGGGATATCC 1171
QY 20 AlaCysLeuAsnProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 1172 TCGTACCAGAGCCACCTACCTGATATGAGTTCGCCGGAAGTGGATGCTGCACAAAC 1231
QY 40 ValLeuHisGlyAlaGluProLeuProTyTyThrAlaAlaHisAlaMetSerTrpCys 59
Db 1232 TTCCTTCACGGC---AAGGTGAAGCGGCATATCTCTGCGCCACGCCCATGAGCTGGCG 1288
QY 60 LeuGlnCysSerGlnGlyValAlaTyLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 1289 CGCCAAATGTGCAGAGGTCTGGCATAITTCATGCCATGACGCCAAACCACTAATACAT 1348
QY 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
Db 1349 CGCGACGTGAAGCGCGTGAACCTCTCTTGACCACAAAGGAGCGCAATCTGAAGATATGC 1408
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
Db 1409 GACTTCGCGCAGGTGGCGGCAAGTCGACCATGATGATGATGATGATGATGATGATGATG 1468
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTyTrSerGluLysCysAspValPheSer 139
Db 1469 TGGATGGCGCCCGAGGTCTTCGAAGGCTCCAAGTATACGGAAGTGTGACATTTTAGC 1528
QY 140 TrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGly 159
Db 1529 TGGGCCATTGTTCTATGGAGGTTCTGTCAGGAAGCAGCCCTTTAAAGGCATCGACAAT 1588
QY 160 ProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsn 179
Db 1589 ---GCCTACCAATCCAGTGAAGATCTACAAGGGTGAACGCCCGCGCTGCTGACCACT 1645

```

```

QY 180 LeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArg 199
Db 1646 TGCCCCAAGCGCATCGAGGACCTGATGACCGCTGCTGGAAACGGTGCCTCGAGGATCGC 1705
QY 200 ProSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyThrPheProGlyAla 219
Db 1706 CCGTCGATGAGTACATAGTGGCGTTATGACGAGATCGTCAAGGACTATACGGGGCG 1765
QY 220 AspGluProLeuGlnTyTr 225
Db 1766 GACAAGGCCCTTGAATAC 1783

RESULT 11
ABL02488
ID ABL02488 standard; cDNA; 10997 BP.
AC ABL02488;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 1946.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB58385.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 1946; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10997 BP; 3237 A; 2362 C; 2382 G; 3016 T; 0 other;

Alignment Scores:
Pred. No.: 3.43e-57 Length: 10997
Score: 599.00 Matches: 130
Percent Similarity: 50.61% Conservative: 35
Best Local Similarity: 39.88% Mismatches: 57
Query Match: 47.84% Indels: 105
DB: 23 Gaps: 5

US-09-830-144-2_COPY_76_303 (1-228) x ABL02488 (1-10997)

```

```

QY      2  G1uLeuAArgGlnLeuSerArgValAsnHisProAsnHisValLeuTyrGly----- 19
DB      2652  GAGGTGAACAGATTGTGTGGCGGTGAACCCGAAACATCATCTCTGACCGCGGATATCC 2711
QY      20  A1AcysLeuAsnProValCysLeuValMetGluTyrAlaGlyGlySerLeuTyrAsn 39
DB      2712  TCGTACCAGCAGGCCACTTACCTGATATAGAGTTGCGGAAGGTGATCGCTGCACAAAC 2771
QY      40  ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
DB      2772  TTCCTTCACGGC---AAGGTGAAGCCGGCATATTCTGTGCCACAGCCATGAGCTGGGGG 2828
QY      60  LeuGlnCysSerGln----- 64
DB      2829  CGCCAAATGTGCAGAGGTAGTGTGTGTGAGGGGGGGGGGTTATTAACGTAGGGCTCA 2888
QY      64  ----- 64
DB      2889  CCCCCGAGCGCTTTCATCCAAATTCGTTAAATTATCTGTGAATAATGACTCAAAATGG 2948
QY      64  ----- 64
DB      2949  ATTGTGTATTACATTCATGCTTGATGATTACTTAACAATTTCAATAGCTATCCGAT 3008
QY      64  ----- 64
DB      3009  AACCAGAATTAATAATAGTAATTTGTTAAAGTTTAAAGTTCATTAATTCATTTG 3068
QY      65  -----GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHisAsr 80
DB      3069  TTTTCTCTCTAGGGGTCTGATTTTGTGATGTCATGACGCAAAACCATATATCATCG 3128
QY      80  gAspleuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysHisCysAs 100
DB      3129  CCACGTGAAGCCGCTGAACCTGCTCTTGAACAAGGAGCCGATCTGAAGATATGCGA 3188
QY      100  pHeGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 120
DB      3189  CTTTGGCAGCGGTGGCGGAGAACTGCAATGATGACCAACAATCCGGCAGTGGCGCTTG 3248
QY      120  pMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTr 140
DB      3249  GATGGCGCCCGAGGTCTTGAAGGCTCCACAGATATCGGAAGTGTGACATTTTACCTG 3308
QY      140  pGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGluIleGlyGlyPr 160
DB      3309  GGCCATTGTTCTATGGAGGTCTGTCCAGGAAGCAGCCCTTTAAAGCATCGACAAAT-- 3366
QY      160  oAlaPheArgIleMetTrpAlaValHisAspGlyThr----- 172
DB      3367  -GCCCTAACCATTCACGTGGAAGATCTTACAAAGGTTGC-GTCTCCCAATTCACCTTTTTC 3424
QY      173  -----ArgProProLeuIleLysAsnLe 180
DB      3425  ATCGAGCTTATCGAGATCTGTGCTCTCGCAGGTGAAGCCCGCGCTGCATCGACACTTG 3484
QY      180  uProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnAsrPr 200
DB      3485  CCCCAGGCGCATCGAGGAGCTGTATGATCCGCTGCGGAAAAAGGTGGCCGAGATCGGCC 3544
QY      200  oSerMetGluGluIleValLysIleMetThrHisLeuMetArgTyrPheProGlyAlaAs 220
DB      3545  GTTCGATGCAGTACATAGTGGGGGTATGCACGAGATCTGTCAAGGACTATATACGGGGCGGA 3604
QY      220  pGluProLeuGlnTyr 225
DB      3605  CAAGGCCCTGGAATAC 3620

```

```

RESULT 12
ABL08337
ID      ABL08337 standard; cDNA; 759 BP.
XX

```

```

AC      ABL08337;
XX
XX      26-MAR-2002 (first entry)
DT
XX
DE      Drosophila melanogaster expressed polynucleotide SEQ ID NO 19493.
KW      Drosophila; developmental biology; cell signalling; insecticide;
KW      pharmaceutical; gene; ss.
XX
OS      Drosophila melanogaster.
XX
XX      MO200171042-A2.
XX
XX      27-SEP-2001.
XX
XX      23-MAR-2001; 2001WO-US09231.
XX
XX      23-MAR-2000; 2000US-191637P.
XX      11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
XX      WPI; 2001-656860/75.
XX      P-PADB; ABB64234.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX      PT interactions -
XX
XX      Claim 1, SEQ ID NO 19493; 21pp + Sequence listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      CC capable of detecting 1000 or more genes from Drosophila. The invention is
XX      CC useful in developmental biology and in elucidating cell signalling and
XX      CC cell-cell interactions in higher eukaryotes for the development of
XX      CC insecticides, therapeutics and pharmaceutical drugs. The invention
XX      CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX      CC sequences (AB101840-AB16175) and the encoded proteins
XX      CC (AB57737-AB57072).
XX      CC The sequence data for this patent did not form part of the printed
XX      CC specification, but was obtained in electronic format directly from WIPO
XX      CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX      SQ Sequence 759 BP; 226 A; 162 C; 198 G; 173 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 1,128-37 Length: 759
XX      Score: 414.00 Matches: 82
XX      Percent Similarity: 59.69% Conservative: 35
XX      Best Local Similarity: 41.84% Mismatches: 53
XX      Query Match: 33.07% Indels: 26
XX      DB: 23 Gaps: 4
XX
XX      US-09-830-144-2_COPY_76_303 (1-228) x ABL08337 (1-759)
XX
QY      2  G1uLeuAArgGlnLeuSerArgValAsnHisProAsnHisValLeuTyrGlyAlaCys 21
DB      160  GAGATCTACAGGTGACAAAGCCGACATGTCACATAGTTGACTACGGCACATCG 219
QY      22  LeuAsnProValCys-----LeuValMetGluTyrAlaGlyGlySerLeuTyrAsn 39
DB      220  AGGACAGAGAGGTGCGCCCTGTGTGATGGAATTCGTAAGCGGTGATCTGTGCCAGT 279
QY      40  ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
DB      280  TTTCTGCAC---GCGAAAGCAAGCCAAAGTTATTCGATGCCAGCCTTCACTGAGGCG 336
QY      60  LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHis 79
DB      337  CATGAGATGCTCAGGGGCACTACCTATCTGATGCGCATGCGACGCAAGCAAGTATTCAT 396

```



```
Qy 80 ArgAspLeuLysProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 397 CCGATATAAGCCCACTCAATACACTGCTATCGGAGAGGACTCAACTGAGATTGC 456
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAla 119
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 457 GATTTCCGAACCTGTTGTGACCTATCCCAATCGATATCGTCAATCGCGGCACCTGCAGA 516
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 120 TrpMetAlaProGlu----- 124
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 517 TACAAGGCCCGCAGGTGAAGGAGTGTGTTGATTCAATCAATCGAATAATATCAAC 576
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 125 -----ValPheGluGlySerAsnTyrSerGluLysCysAspVal 137
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 577 CAACCAACCGGCTTCAAAAGTTCTACAAGGAATAAACCCGATGAAAGTGGATGTG 636
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 138 PheSerTrpGlyIleIleLeuTrpGluValIleThrArgLysProPheAspGluIle 157
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 637 TATAGTTGGCTATTACCTTTTGGGAATATTGTGCGCAAGGAGCCATTGAGCAATAT 696
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 158 GlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArg 173
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 697 ---AATACGCTTTTGAACCTATACATGGCTATTATTAATGAAGGCAAGAGA 741
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 13
ABL70018
ID ABL70018 standard; DNA; 3454 BP.
XX
AC ABL70018;
XX
DT 15-MAY-2002 (first entry)
XX
DE Pancreas cancer related gene sequence SEQ ID NO:8355.
XX
KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;
KW gene; ds.
XX
OS Homo sapiens.
XX
PN WO200194629-A2.
XX
PD 13-DEC-2001.
XX
PF 30-MAY-2001; 2001WO-US10838.
XX
PR 05-JUN-2000; 2000US-209473P.
PR 05-JUN-2000; 2000US-209531P.
PR 18-SEP-2000; 2000US-231333P.
PR 18-SEP-2000; 2000US-233617P.
PR 20-SEP-2000; 2000US-234009P.
PR 20-SEP-2000; 2000US-234034P.
PR 20-SEP-2000; 2000US-234052P.
PR 22-SEP-2000; 2000US-234509P.
PR 22-SEP-2000; 2000US-234567P.
PR 25-SEP-2000; 2000US-234923P.
PR 25-SEP-2000; 2000US-234924P.
PR 25-SEP-2000; 2000US-235077P.
PR 25-SEP-2000; 2000US-235082P.
PR 25-SEP-2000; 2000US-235134P.
PR 25-SEP-2000; 2000US-235280P.
PR 26-SEP-2000; 2000US-235637P.
PR 26-SEP-2000; 2000US-235638P.
PR 27-SEP-2000; 2000US-235711P.
PR 27-SEP-2000; 2000US-235720P.
PR 27-SEP-2000; 2000US-235840P.
PR 27-SEP-2000; 2000US-235863P.
PR 28-SEP-2000; 2000US-236028P.
PR 28-SEP-2000; 2000US-236032P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236034P.
PR 28-SEP-2000; 2000US-236109P.
```

```
PR 28-SEP-2000; 2000US-236111P.
PR 29-SEP-2000; 2000US-236842P.
PR 29-SEP-2000; 2000US-236891P.
PR 02-OCT-2000; 2000US-237172P.
PR 02-OCT-2000; 2000US-237173P.
PR 02-OCT-2000; 2000US-237278P.
PR 02-OCT-2000; 2000US-237294P.
PR 02-OCT-2000; 2000US-237295P.
PR 02-OCT-2000; 2000US-237316P.
PR 03-OCT-2000; 2000US-237425P.
PR 03-OCT-2000; 2000US-237598P.
PR 03-OCT-2000; 2000US-237604P.
PR 03-OCT-2000; 2000US-237606P.
PR 03-OCT-2000; 2000US-237608P.
PR 01-NOV-2000; 2000US-244867P.
PR 01-NOV-2000; 2000US-245084P.
XX
XX (AVAL-) AVALON PHARM.
XX
XX Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;
PI Soppet DR, Weaver Z;
XX
XX WPI; 2002-188264/24.
XX
XX Screening for anti-neoplastic agent involves exposing cells to a
PT chemical agent to be tested for anti-neoplastic activity, and
PT determining a change in expression of a gene of a signature gene set -
XX
XX Claim 1; SEQ ID 8355; 44pp; English.
XX
XX The present invention describes a method (M1) for screening for an
CC anti-neoplastic agent. The method involves exposing cells to a chemical
CC agent to be tested for anti-neoplastic activity, determining a change in
CC expression of at least one gene (I) of a signature gene set, where (I)
CC comprises a sequence (S) selected from 8447 sequences (given in ABL61664
CC to ABL70010), or is at least 95% identical to (S), where a change in
CC expression is indicative of anti-neoplastic activity. (I) has cycostatic
CC activity and can be used in gene therapy. M1 can be used for screening
CC an anti-neoplastic agent, and can be used for producing a product which
CC is the data collected with respect to the anti-neoplastic agent as a
CC result of M1, and the data is sufficient to convey the chemical
CC structure and/or properties of the agent. M1 can be used in the
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine
CC carcinoma, papillary carcinoma and Wilm's tumour.
XX
XX SQ Sequence 3454 BP; 594 A; 1217 C; 1136 G; 507 T; 0 other;
XX
XX Alignment Scores:
Pred. No.: 2,79e-34 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 24 Gaps: 6
XX
XX US-09-830-144-2_COPY_76_303 (1-228) x ABL70018 (1-3454)
Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 721 GAAGCCCGCTCTTTGGAGCCCTGCGACCCCAACATAATTGCCCTTAGGGGCGCTGC 780
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 781 CTCACCCCCCACCTCTGCTAGTGATGGAGTATGCCGGGTGGTGGCTAGGAGCAGG 840
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCys 59
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 841 GTGCTGGCAGGTGCGCGGGTGCCACCTCACGTG-----CTGTCACTGGGCT 888
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
Db      889  GTCCAGGTGCCCCGGGCGCATGAACCTACCTACCAATGATGCCCTTGCCCATCATCCAC 948
Qy      80   ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db      949  CCGGACCTCAAGTCCATCAACATCTGATCTCGAGGCCATGAGAACCAACCTCGCA 1008
Qy      93   GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---IleGlnThrHisMet 111
Db      1009 GACACGCTGCTCAAGATCAACGCGACTTCGGCTCGCCCGGAGTGGCACAAGACCAACG 1068
Qy      112  ThrAspAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluLysSerAsnTyr 131
Db      1069 ATGAGCGCTCGCGGGAGCACTACCGCTGATGCCCGCGAGATTATCCGTCTCCCTTC 1128
Qy      132  SerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArg 151
Db      1129  TCCAAAGCAGTGTATGTCTGAGCTTCGGGGTGCTGCTGGAGAGCTGTACGGGGGAG 1188
Qy      152  LysProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db      1189 GTCCCTTACCGTAGATC-----GAGGCTTGGCCGCTGATGGCGGTATGAT 1242
Qy      171  GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 190
Db      1243 AAGCTAGCGCTGCCCATTCCTCCACAGTGCCTCCGAGCCCTTGCCCGCTCTCGAGGAA 1302
Qy      191  CysTrpSerLysAspProSerGlnArgProSerMetGluIleValLys 207
Db      1303 TCTGTGGACCCAGACCCCGCGGGCCAGATTTCGTAGCATCTTGAG 1353

RESULT 14
AAC43254
ID      AAC43254 standard: DNA; 1224 BP.
XX      AAC43254;
XX      17-OCT-2000 (first entry)
DT      XX
DE      Arabidopsis thaliana DNA fragment SEQ ID NO: 38594.
XX      XX
XX      Hybridisation assay; genetic mapping; gene expression control;
KW      protein identification; signal transduction pathway;
KW      metabolic pathway; promoter; termination sequence; ss.
XX      OS
XX      Arabidopsis thaliana.
XX      EP1033405-A2.
XX      PD
XX      06-SEP-2000.
XX      PF
XX      25-FEB-2000; 2000EP-0301439.
XX      XX
XX      25-FEB-1999; 99US-0121825.
XX      PR      05-MAR-1999; 99US-0123180.
XX      PR      09-MAR-1999; 99US-0123548.
XX      PR      23-MAR-1999; 99US-0125788.
XX      PR      25-MAR-1999; 99US-0126264.
XX      PR      29-MAR-1999; 99US-0126785.
XX      PR      01-APR-1999; 99US-0127462.
XX      PR      06-APR-1999; 99US-0128234.
XX      PR      08-APR-1999; 99US-0128714.
XX      PR      16-APR-1999; 99US-0129845.
XX      PR      19-APR-1999; 99US-0130077.
XX      PR      21-APR-1999; 99US-0130449.
XX      PR      23-APR-1999; 99US-0130510.
XX      PR      23-APR-1999; 99US-0130891.
XX      PR      28-APR-1999; 99US-0131449.
XX      PR      30-APR-1999; 99US-0132048.
XX      PR      30-APR-1999; 99US-0132407.
XX      PR      04-MAY-1999; 99US-0132484.
XX      PR      05-MAY-1999; 99US-0132485.
XX      PR      06-MAY-1999; 99US-0132486.
```

```
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-0134221.
PR      14-MAY-1999; 99US-0134370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-0134941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135324.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136392.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137528.
PR      04-JUN-1999; 99US-0137502.
PR      07-JUN-1999; 99US-0137724.
PR      08-JUN-1999; 99US-0138094.
PR      10-JUN-1999; 99US-0138540.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      17-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.
PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139763.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140991.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      19-JUL-1999; 99US-0144331.
PR      19-JUL-1999; 99US-0144333.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144334.
PR      19-JUL-1999; 99US-0144335.
PR      20-JUL-1999; 99US-0144332.
PR      20-JUL-1999; 99US-0144632.
PR      21-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0144884.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145085.
PR      22-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145089.
```


AC AAC39537;
XX 17-OCT-2000 (first entry)
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 24996.
DE
XX
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131445.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139451.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145214.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145813.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:59:01 ; Search time 2220 Seconds
(without alignments)
1663.320 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303
Perfect score: 1252
Sequence: 1 VELRQLSRVNHNPVIVKLYG.....MTHLMRYFPGADEPLQYPCQ 228

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlp
-DB/cgn2_1/USPTO.spool/US09830144/runat_04122002_141754_4354/app_query.fasta_1.391
-QS=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG_1_1_1716 @runat_04122002_141754_4354 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:.*
1: em_estba:.*
2: em_esthum:.*
3: em_estin:.*
4: em_estmu:.*
5: em_estov:.*
6: em_estpl:.*
7: em_estro:.*
8: em_htc:.*
9: gb_est1:.*
10: gb_est2:.*
11: gb_htc:.*
12: gb_est3:.*
13: gb_est4:.*
14: gb_est5:.*
15: em_estfun:.*
16: em_estom:.*
17: gb_gss:.*
18: em_gss_hum:.*
19: em_gss_inv:.*
20: em_gss_pln:.*
21: em_gss_vrt:.*
22: em_gss_fun:.*
23: em_gss_mam:.*
24: em_gss_mus:.*
25: em_gss_other:.*
26: em_gss_pro:.*
27: em_gss_rod:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1167	93.2	739	13	BI696710	BI696710 603347214
2	1069	85.4	604	13	BM426610	BM426610 pfg2n.pko
3	1016	81.2	1062	13	BM554120	BM554120 AGENCOURT
4	984	78.6	646	13	BJ073883	BJ073883 BJ073883
5	982	78.4	688	13	BJ062988	BJ062988 BJ062988
6	982	78.4	696	13	BJ074867	BJ074867 BJ074867
7	953.5	76.2	910	12	BF780358	BF780358 602103276
8	932.5	74.5	998	9	AL525728	AL525728 AL525728
9	847	67.7	542	10	AW960377	AW960377 EST372448
10	823	65.7	641	14	BQ386875	BQ386875 NISC_mn20
11	787	62.9	827	9	AL550589	AL550589 AL550589
12	753	60.1	1006	14	BQ219348	BQ219348 AGENCOURT
13	748	59.7	686	12	BG548917	BG548917 602575750
14	741	59.2	771	13	BI093821	BI093821 602860601
15	737	58.9	682	10	AV398933	AV398933 AV398933
16	715	57.1	586	13	BJ035129	BJ035129 BJ035129
17	707	56.5	681	10	AV398935	AV398935 AV398935
18	705	56.3	929	9	AL520975	AL520975 AL520975
19	631.5	50.4	1054	14	BQ067693	BQ067693 AGENCOURT
20	631	50.4	607	9	AI541706	AI541706 SD06739.5
21	628	50.2	594	13	BJ014293	BJ014293 BJ014293
22	587	46.9	747	13	BI334774	BI334774 602999677
23	559	44.6	516	13	BJ006787	BJ006787 BJ006787
24	557	44.5	858	13	BG916821	BG916821 602816079
25	553	44.2	711	10	AV718168	AV718168 AV718168
26	549.5	43.9	760	12	BG773488	BG773488 602720209
27	549	43.8	822	13	BI757369	BI757369 603029591
28	548	43.8	538	14	BQ557823	BQ557823 H4048E02-
29	540	43.1	745	13	BI767913	BI767913 603061057
30	515.5	41.2	647	13	BJ008789	BJ008789 BJ008789
31	497	39.7	599	10	BE284386	BE284386 601087550
32	496	39.6	286	9	AA450856	AA450856 VQ55H03.1
33	483	38.6	585	9	AL773651	AL773651 AL773651
34	474	37.9	718	9	AL652289	AL652289 AL652289
35	469.5	37.5	1191	13	BM472107	BM472107 AGENCOURT
36	457	36.5	1033	12	BE882182	BE882182 601505283
37	452	36.1	647	10	BE665645	BE665645 BB665645
38	445	35.5	569	9	AU177912	AU177912 AU177912
39	443	35.4	338	12	BE707048	BE707048 MRO-HT040
40	438	35.0	831	14	BQ736206	BQ736206 AGENCOURT
41	437.5	34.9	845	12	BF027003	BF027003 601671216
42	430	34.3	941	12	BG029948	BG029948 602297101
43	420	33.5	248	9	AI614762	AI614762 VQ55H03-Y
44	387.5	31.0	689	14	BQ745060	BQ745060 UI-M-EHOP
45	374	29.9	740	13	BJ441629	BJ441629 BJ441629

ALIGNMENTS

RESULT 1
BI696710
LOCUS 603347214F1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5374971 5',
DEFINITION 739 bp mRNA linear EST 18-SEP-2001
mRNA sequence.
ACCESSION BI696710
VERSION BI696710.1 GI:15659339
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 739)
NIH-MGC <http://mgs.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 DNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LNLN)
 DNA distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.llnl.gov
 Plate: LNLN1954 row: k column: 04
 High quality sequence stifiers: 717.
 Location/Qualifiers
 1. 739
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5374971"
 /clone_lib="NCI CGAP Mam2"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: PCMV-SPORT6; Site 1: SalI;
 Site 2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"
 BASE COUNT 195 a 169 c 188 g 187 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,62e-124 Length: 739
 Score: 1167.00 Matches: 220
 Percent Similarity: 98.65% Conservative: 0
 Best Local Similarity: 98.65% Mismatches: 1
 Query Match: 93.21% Indels: 3
 DB: 13 Gaps: 0
 US-09-830-144-2_copy_76_303 (1-228) x B1696710 (1-739)
 QY 8 ArgValAlaenHisPioAsnIleValIysLeuTYrGlyAlaCysLeuAsnProValCysLeu 27
 Db 3 CGTGTGAACCATCTTAACATTTGTCAAGTTGTACGAGCGCTGCTGAATCCAGTATGTCTT 62
 QY 28 ValMetGluTYrAlaGluGlyGlySerLeuTYrAsnValIleuHisGlyAlaGluProleu 47
 Db 63 GTGATGAATATGACAGAGGGGGCTCATTTGTAATATGCTGCATGCTGACCAACATTTG 122
 QY 48 ProTYrTYrThAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGly-ValAl 67
 Db 123 CCTTACTACACTGCTGCTCATCCCATGAGCTGCTTTTACAGTGTCCCAAGTATGTGC 182
 QY 67 aTYrIleuHisSerMetGlnProIysAlaLeuIleHisArgAspLeuLysProProAsnIle 87
 Db 183 TTACCTGCACAGCATGCGCCCAAGCGCTGATTCACAGGGAGCCTCAAGCCTCCAACTT 242
 QY 87 ILeuLeuValAlaGlyGlyThrValIleuLysIleCysAspPheGlyThrAlaCysAspI 107
 Db 243 GCTGCTGGTGGAGAGGAGGAGCTTCTAAATCTTGCATTTTGTATACAGCTTGTGACAT 302
 QY 107 eGlnThrHisMetTrpAsnAsnLysGlySerAlaIleTrpMetAlaProGluValIlePheG 127
 Db 303 CCAACACACATGACCATTAATAAGGAGAGTCTGCTTGATGGGCGCTGAAGTCTTTGA 362
 QY 127 uGlySerAsnTYrSerGlyLysCysAspValIlePheSerTrpGlyIleIleLeuTrpGluVal 147
 Db 363 AGGTGTGCAATTAAGTGAAGTGTGATGTCTTCAGCTGGGGTATTATCTCTGGGAAGT 422
 QY 147 ILeuThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAl 167
 Db 423 GATTAACAGCCCGAATCCCTTGTGATGAGATCGGTGGCCAGCTTTCAGAAATCATGTGGGC 482
 QY 167 aValHisAsnGlyThrArgProProIleuIleLysAsnLeuProLysProIleGluSerIle 187

Db 483 TGTTCATATGCGACCTGCGACCACTGATCAAAAATTTACTAAGCCCATTTGAGAGCTT 542
 QY 187 uMetThrArgCys-TrpSerLysAspProSerGlnArgProSerMetGluGluIleValI 207
 Db 543 GATGACACGCTGTGTGTGTTCAAGACCAATCTCAGGC-CCTTAATGAGGAATTTGTA 601
 QY 207 yAlleMetThrHisLeuMetArgTYrPheProGlyValAspGluProLeuGlnTYrProC 227
 Db 602 AAATTAATGACTCTGATGATGCGGTACTTCCAGAGCGGATGAGCCATTAAGTATCTT 661
 QY 227 yGcIn 228
 Db 662 GTCAAG 666
 RESULT 2
 BM426610
 LOCUS
 DEFINITION pgf2n.pk004.d15 Normalized Chicken Abdominal Fat Library (pgf2n)
 GALLUS gallus CDNA clone pgf2n.pk004.d15 5' similar to
 emb|CAB87605.1 (AL121964) d015461.1.4 (mitogen-activated protein
 kinase kinase kinase 7 (TGF-beta activated kinase 1d (TAK1))) [Homo
 sapiens], mRNA sequence.
 ACCESSION BM426610
 VERSION 1
 KEYWORDS EST.
 SOURCE chicken.
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 604)
 Cogburn,L.A., Morgan,R. and Burnside,J.
 ESTs from Normalized Chicken fat CDNA library-USDA/IFAFS Animal
 Genome Project
 Unpublished (2002)
 JOURNAL Contact: Larry A. Cogburn
 University of Delaware
 Townsend Hall, Newark, DE 19717, USA
 Tel: 302-831-1335
 Fax: 302-831-2822
 Email: cogburnudel.edu, www.chickest.udel.edu.
 Location/Qualifiers
 source
 1. 604
 /organism="Gallus gallus"
 /strain="Commercial broiler, Ottawa Research Centre,
 leghorn"
 /db_xref="taxon:9031"
 /clone="pgf2n.pk004.d15"
 /clone_lib="Normalized Chicken Abdominal Fat Library
 (pgf2n)"
 /sex="Male and Female"
 /tissue_type="Abdominal Fat"
 /dev_stage="Embryonic (d18,d19); post-hatch (d1,w3,w7,w9
 w16,1yr)"
 /lab_host="E. coli EMDH10B"
 /note="Vector: PCMVSPORT6; Library made from equivalent
 pools of total RNA isolated from each developmental age
 (across strains); Single pass sequencing from 5'-end"
 BASE COUNT 167 a 125 c 140 g 167 t 5 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.6e-113 Length: 604
 Score: 1069.00 Matches: 194
 Percent Similarity: 98.99% Conservative: 2
 Best Local Similarity: 97.98% Mismatches: 2
 Query Match: 85.38% Indels: 0
 DB: 13 Gaps: 0
 US-09-830-144-2_copy_76_303 (1-228) x BM426610 (1-604)
 QY 30 GluTYrAlaGluGlyGlySerLeuTYrAsnValIleuHisGlyAlaGluProleuProTYr 49

Db 2 GAGTATGCTGAGGAGTTCTCTGTACAAATGCTGTCATGCTGCTGAACCTCTCTCCCTCAT 61

Qy 50 TTTTThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGlnGlyValAlaTyrLeu 69

Db 62 TATATCTGCTGCACAGCAATGAGTTGGTTTACAGTGTTCCTCCAGAGTGGCAATCTT 121

Qy 70 HisSerMetGlnProLysAlaAlaLeuHisArgAspLeuLysProProAsnLeuLeu 89

Db 122 CACAGTATGAACCAACCAAGCCCTAATTCACAGAGACCTGAAACACCAAAATTCCTCTTG 181

Qy 90 ValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAlaCysAspIleGlnThr 109

Db 182 GTAGCTGGGGGACAGTTCTTAAAGATCTGTGATTTGGTACAGCTGTGATATCAACA 241

Qy 110 HisMetThrAsnAsnLysGlySerAlaAlaTTPMetAlaProGluValPheGluGlySer 129

Db 242 CACATGACCAACAATPAGGGAAGTCTGCTGGATGGACCTGAAGTTTGGAGGTAGC 301

Qy 130 AsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeuLeuTrpGluValIleThr 149

Db 302 AATTACAGTGAACCAATGTGACGTTTTCAGTTGGGGTATATTCCTTTGGGAGGTAATCACC 361

Qy 150 ArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHis 169

Db 362 CGTAGGAAACCTTTTGTATGAGATTTGGTGTCCAGCTTTCCGCAATAATGTGGGCACTTCAC 421

Qy 170 AsnGlyThrArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThr 189

Db 422 AATGGTACTCGACCACTGATCAAAACTTACCTTAACCAATTTGAGAGTTAATGACC 481

Qy 190 ArgCysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLysIleMet 209

Db 482 CGCTGTTGGTCAAGGATCCCTCACACGACCTTCCATGGAGGAAATTTGTTAAATAATG 541

Qy 210 ThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGlnTyrProCys 227

Db 542 ACACACTTGTATCGGTACTTTCNNNAGCTGATGAACCTCTCGCATATCCCTTGC 595

RESULT 3

BM554120 1062 bp mRNA linear EST 20-FEB-2002

LOCUS AGENCOURT_6546903 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5742353

DEFINITION 5', mRNA sequence.

ACCESSION BM554120

VERSION BM554120.1 GI:18793446

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1. (bases 1 to 1062)

AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgabs-re@mail.nih.gov

Tissue procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Plate: LLNL2760 row: f column: 18

High quality sequence stop: 730.

Location/Qualifiers

1. .1062

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5742353"

/clone_lib="NIH_MGC_119"

/tissue_type="medulla"

/lab_host="DH10B"

FEATURES

source

RESULT 4

BM554120 646 bp mRNA linear EST 11-DEC-2001

LOCUS BU073883

DEFINITION BU073883 NIBB Mochli normalized Xenopus tailbud library Xenopus laevis cDNA clone XL101n05 5', mRNA sequence.

ACCESSION BU073883

VERSION BU073883.1 GI:17504072

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source normal medulla from anonymous male age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.3 kb, insert size range 0.9-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 013. Note: this is a NIH MGC Library."

BASE COUNT 253 a 243 c 298 g 266 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 8.56e-107 Length: 1062

Score: 1016.00 Matches: 203

Percent Similarity: 94.52% Conservativeness: 4

Best Local Similarity: 92.69% Mismatches: 5

Query Match: 81.15% Indels: 7

DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BM554120 (1-1062)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20

Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCTCTAATATTGTAAGCTTTATGAGGCC 467

Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40

Db 468 TGCTTGAATCCAGTGTCTTTGTGATGGAATATGCTGAAGGGGCTCTTTATATAATGTG 527

Qy 41 LeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60

Db 528 CTGCATGGTCTGAACCAATTCGCATATTATCTTACAGCATGCAACCCCAAGCGCTAATTCACAGG 587

Qy 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80

Db 588 CAGTGTTCCTCAAGAGTGGCTTATCTTTCACAGCATGCAACCCCAAGCGCTAATTCACAGG 647

Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAsp 100

Db 648 GACCTGAACCAACCAACTTACTGCTGGTTGCAAGGGGACAGCTTCTAAAAATTTGTGAT 707

Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120

Db 708 TTTGGTACAGCTGTGACATTCACACACATGACCAATTAACCAAGGGGAGTGTCTGCTGG 767

Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140

Db 768 ATGCACCTGAAGTTTTTGAAGGTAGTAATTCAGTGAATAATGACGTTCTTCAGCTGG 827

Qy 141 GlyIleLeuLeuTrpGluValIleThrArgArg-LysProPheAspGluIleGly-P 160

Db 828 GGTATTATTCTTTGGGAAGTGATAACCGTCGGAACCAACCTTTGATGAGATGGTGGCC 887

Qy 160 roAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeu-IleLysAsn 179

Db 888 CAGCTTTCCGAATCATGTGGCTGGTCAATGCTACTCGACCACTGTATAAAAAAT 947

Qy 180 LeuProLysProIleGlu-SerLeuMetThr-ArgCysTrpSerLysAspPro-SerGln 198

Db 948 TTACTAGGCCCATTAAGAACCTGATGATCCCTGTTGGTGGTAAAGATCTTTCCAG 1007

Qy 199 ArgProSerMetGlu-GluIleValLysIleMetThrHisLeu 212

Db 1008 CGCCCTTTCATGGAGGGAATGTTGAAAAATAATGACCTCACTTG 1050

KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 646)
AUTHORS Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.

TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. 646
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL101n05"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute). "

BASE COUNT 190 a 132 c 143 g 181 t
ORIGIN

Alignment Scores:
Pred. No.: 1,99e-103 Length: 646
Score: 984.00 Matches: 182
Percent Similarity: 97.37% Conservative: 3
Best Local Similarity: 95.79% Mismatches: 5
Query Match: 78.59% Indels: 1
Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BU073883 (1-646)

QY 39 AsnValleuHISGLYAlaGluProLeuProTyTYrThrAlaAlaHISAlaMetSerTrp 58
Db 4 AATGTTTGCATGAGAGCTGAACCTTGGCTTACAT-ACGCTGCCCATGCAATGAGTTGG 62
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrlleuHISerMetGlnProLysAlaLeuIle 78
Db 63 TGTTTACAAATGTCGCCAAGAGGTTGCATATTTACATACATGAAGCCAAAGGCTCTGATT 122
QY 79 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 98
Db 123 CACAGGAGCCTCAAAACCAAACTTGTGCGGTAGAGTGAGGACGTTCTTAAGATT 182
QY 99 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 118
Db 183 TGTGACTTTGGTACAGCTGTGATTTCAAGACTCACATGACTAATAACAAAGAAAGTGCA 242
QY 119 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrlserGlnLysCysAspValPhe 138
Db 243 GCATGAGTGGCTCCAGAACTTTTGAAGTAGCAACTACAGCAAAATAATGTGACGTGTTT 302
QY 139 SerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 158
Db 303 AGTTGGGCAATTATCTTTGGGAAGTATATAACCGAATAAACTTTGAGAAATTGGT 362
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys 178
Db 363 GGTCCAGGCTTCGTATATATGTGGGCTGTTCCAAATAGTACTCGGCACCACTAATTAAA 422

QY 179 AsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln 198
Db 423 AATTGGCTTAAGCCTATTGAAAGCTTAATGACTCGCTGCTGCCAAGATCCCCACAA 482
QY 199 ArgProSerMetGluGluIleValLysIleMetThrHisleuMetArgTyrlPheProGly 218
Db 483 AGACCTTCAAATGAGAGGATTTGCAAGATATACACATCTAAGACAGATATTTCTTGG 542
QY 219 AlaAspGluProLeuGlnTyrlProCysGln 228
Db 543 GCAGACGTTCTTACAGTATCTCTTGTGAG 572

RESULT 5
LOCUS BU062988 688 bp mRNA linear EST 10-DEC-2001
DEFINITION BU062988 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL069m17 5', mRNA sequence.
ACCESSION BU062988
VERSION BU062988.1 GI:17470746
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 688)
AUTHORS Kiteyama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
Y.

TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadaeu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. 688
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL069m17"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Nieuwkoop and Faber. Library is subtracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute). "

BASE COUNT 206 a 139 c 154 g 188 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 3,74e-103 Length: 688
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 78.43% Indels: 0
Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BU062988 (1-688)

QY 45 GluProLeuProTyTYrThrAlaAlaHISAlaMetSerTrpCysLeuGlnCysSerGln 64
Db 22 GAACCTTTCCTTACATATCTGCTGCCATGCAAGAGTGGTGTTTCAATGTGCCCA 81
QY 65 GlyValAlaTyrlleuHISerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTTACATGATGAAGCCAAAGCTGTGATTCACAGGAGCCTCAAAACA 141
QY 85 ProAsnLeuLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 104

```

Db 142 CCAAACTTGTGTGCTAGCTGGAGCACTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 124
Db 202 TGTGATATTACAGACTCACATGACTAATAACAAGGAAGTGCAGCATGGATGCTCCAGAA 261
Qy 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeu 144
Db 262 GTTTTGAAGGTAGCAACTACACGGGAAAATGTGACGTGTTAGTTGGGCATTATTCTT 321
Qy 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAAGAAACCTTTCGATGAAATGGTGGTCCAGCGTTCGTGATA 381
Qy 165 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 184
Db 382 ATGTGGGCTGTTACAAATGGTACTCGGCACCATTAATTAATAAATTTGCTTAAGGCTATT 441
Qy 185 GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 442 GAAAGCTTAATGACTCGCTGCTGCTCCAAAGATCCCCCAAGACCTTCAATGAGGAG 501
Qy 205 IleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 224
Db 502 ATTGTCAAGATAATGACATCTAAAGCAGTATTTTCTCGGAGCAGACGTTTCTTACAG 561
Qy 225 TyrProCysGln 228
Db 562 TATCCTTGTCAG 573

RESULT 6
BJ074867
LOCUS
DEFINITION
  BJ074867 NTBB Mochii normalized Xenopus tailbud library Xenopus
  laevis cDNA clone XL071110 5', mRNA sequence.
ACCESSION
  BJ074867.1 GI:17505056
VERSION
  EST.
KEYWORDS
  African clawed frog.
SOURCE
  Xenopus laevis
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
  Xenopodinae; Xenopus.
REFERENCE
  1 (bases 1 to 696)
  Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
  Y.
TITLE
  Expressed genes in X. laevis embryo
JOURNAL
  Unpublished (2001)
COMMENT
  Contact: Tadasu Shin-i
  Center For Genetic Resource Information
  National Institute of Genetics
  1111 Yata, Mishima, Shizuoka 411-8540, Japan
  Tel: 81-559-81-6856
  Fax: 81-559-81-6855
  Email: tshini@genes.nig.ac.jp.
FEATURES
  Location/Qualifiers
    1..696
      /organism="Xenopus laevis"
      /db_xref="taxon:8355"
      /clone="XL071110"
      /clone_lib="NIBB Mochii normalized Xenopus tailbud
      library"
      /tissue_type="whole embryo"
      /dev_stage="stage 25"
      /notes="Vector: pBSRN3; Site.1: NotI; Site.2: EcoRI; cDNAs
      were oligo-dT primed and directionally cloned. Staging
      according to Nieuwkoop and Faber. Library is substracted
      and was constructed by N. Garrett and A.M. Zorn,
      (Wellcome/CRC Institute)."
BASE COUNT
  210 a 142 c 155 g 188 t 1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3,81e-103 Length: 696
Score: 982.00 Matches: 177
Percent Similarity: 97.83% Conservative: 3
Best Local Similarity: 96.20% Mismatches: 4
Query Match: 78.43% Indels: 0
DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BJ074867 (1-696)

Qy 45 GluProLeuProTyrTrpThrAlaAlaHisAlaMetSerTrpCysLeuGlnCysSerGln 64
Db 22 GAACCTTTTGCTTACTACTGCTGCCATGCAATGAGTTGGTGTTTACAAATGTGCCAA 81
Qy 65 GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArgAspLeuLysPro 84
Db 82 GGAGTTGCATATTTCATATAGCATGAAGCCAAAGGCTCTGATTACAGGGACCTCAACCA 141
Qy 85 ProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAspPheGlyThrAla 104
Db 142 CCAAACTTGTGTGCTAGCTGGAGGCACTGTTCTTAAGATTGTGACTTTGGTACAGCC 201
Qy 105 CysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGlu 124
Db 202 TGTGATATTACAGACTCACATGACTAATAACAAGGAAGTGCAGCATGGATGCTCCAGAA 261
Qy 125 ValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrpGlyIleLeu 144
Db 262 GTTTTGAAGGTAGCAACTACACGGGAAAATGTGACGTGTTAGTTGGGCATTATTCTT 321
Qy 145 TrpGluValIleThrArgArgLysProPheAspGluIleGlyGlyProAlaPheArgIle 164
Db 322 TGGGAAGTAATAACCCGAAGAAACCTTTCGATGAAATGGTGGTCCAGCGTTCGTGATA 381
Qy 165 MetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeuProLysProIle 184
Db 382 ATGTGGGCTGTTCAAAATGGTACTCGGCACCATTAATTAATAAATTTGCTTAAGGCTATT 441
Qy 185 GluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgProSerMetGluGlu 204
Db 442 GAAAGCTTAATGACTCGCTGCTGCTCCAAAGATCCCCCAAGACCTTCAATGAGGAG 501
Qy 205 IleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAspGluProLeuGln 224
Db 502 ATTGTCAAGATAATGACATCTAAAGCAGTATTTTCTCGGAGCAGACGTTTCTTACAG 561
Qy 225 TyrProCysGln 228
Db 562 TATCCTTGTCAG 573

RESULT 7
BJ0780358
LOCUS
DEFINITION
  602103276F1 NCI_CGAP_Kid14 Mus musculus cDNA clone IMAGE:4221379
  5', mRNA sequence.
ACCESSION
  BJ0780358
VERSION
  BJ0780358.1 GI:12085481
KEYWORDS
  EST.
SOURCE
  house mouse.
  ORGANISM
    Mus musculus
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
  1 (bases 1 to 910)
  NIH-MGC http://mgs.nci.nih.gov/
  National Institutes of Health, Mammalian Gene Collection (MGC)
  UNPUBLISHED (1999)
  TITLE
  Unpublished (1999)
  AUTHORS
  Contact: Robert Strausberg, Ph.D.
  Email: cgabs-remail.nih.gov
  Tissue Procurement: Jeffrey B. Green, M.D.
  cDNA Library Preparation: Life Technologies, Inc.
  DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
  DNA Sequencing by: Incyte Genomics, Inc.
  Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/BLML at:

http://image.jlml.gov
Plate: LLM9806 row: h column: 20
High quality sequence stop: 664.
Location/Qualifiers

FEATURES

1..910

/organism="Mus musculus"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone_image="4221379"

/clone_lib="NCI CGAP Kid4"

/lab_host="DH10B (TI phage-resistant)"

/note="Organ: kidney; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt."

Average insert size 1.75 kb. Constructed by Life Technologies. Note: this is a NCI CGAP Library. |"

BASE COUNT 247 a 209 c 246 g 207 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 1.15e-99 Length: 910
Score: 953.50 Matches: 194
Percent Similarity: 92.86% Conservative: 1
Best Local Similarity: 92.38% Mismatches: 6
Query Match: 76.16% Indels: 12
Gaps: 1

US-09-830-144-2_COPY_76_303 (1-228) x BF780358 (1-910)

```

QY 28 ValmetGluTyrAlaGluGlySer-Leu-----TyrAsnVa 40
DB 4 GTGATGGAATATGACAGAGGGGCTCATGATATATGTTGTCGCTTCTTCAAGT 63
QY 40 lleuhsGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTyrCysLe 60
DB 64 GCTGATGCTGCTGAACCATCTGCTTACTACACTGCTCTCATGCTGAGCTGTTT 123
QY 60 uGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProGlyAlaLeuHisAr 80
DB 124 AAGAGTTCCTCCAGAGAGGCTTACCTGACAGCATGAGCCCAAGGCTGATTCACAG 183
QY 80 GAspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCysAs 100
DB 164 GACCTCAAGCTCCAACTGCTGCTGTCAGAGAGGACAGTCTTAAAAATTCGCA 243
QY 100 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaIaTr 120
DB 244 TTTTGTGACAGCTGTGACATCAACAACACATGACCAATATATAAGGAGTGTGCTTG 303
QY 120 pMetAlaProGluValPheGlyLysSerAsnTyrSerGlyLysCysAspValPheSerTr 140
DB 304 GATGGCGCTGAAGTGTGAAGTAGAATTAACATTAAGTGAAGTGTGCTTCAAGCTG 363
QY 140 pGlyIleLeuLeuTyrGluValIle-ThrArgArgLysProPheAspGluIleGlyLP 160
DB 364 GGGTATTAATCCTCTGGAAGTATATAACCGCGAAACCTTCATGATGATCGTGCC 423
QY 160 roAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeuLysAsnL 180
DB 424 CAGCTTTCAGATCATGTGGCTGTTCATTAATGACATCGACCACTGATCAAAAATT 483
QY 180 euProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProserGlnArgP 200
DB 484 TACCTAAGCCATTAAGAGCTTGATGACACGGTG-TGGTCTAAGAGACCATCTAGCG-C 541
QY 200 roSerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGlyAla 220
DB 542 CTTCAATGAGAGAAATTGTGAATAATAGACTCACTTATCGGTAATCTCCAGAGACGG 601
QY 220 spGluProLeuGlnTyrProCysGln 228
DB 602 ATGAGCA-TTACAGTATACTGTGTGAG 626

```

RESULT 8

AL525728 998 bp mRNA linear EST 13-FEB-2001

LOCUS

AL525728 LIT NFL003 NBC3 Homo sapiens cDNA clone CS0DC013YE20 5

DEFINITION

prime, mRNA sequence.

ACCESSION

AL525728 GI:12789221

VERSION

EST.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

AUTHORS

1 (bases 1 to 998)

TITLE

Full-length cDNA libraries and normalization

JOURNAL

Unpublished (2001)

COMMENT

Genoscope - Centre National de Sequencage

FEATURES

BP 191 91006 EVRY cedex - France

source

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

Location/Qualifiers

1..998

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DC013YE20"

/clone_lib="LIT_NFL003_NBC3"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed

by Life Technologies. Contact: Feng Liang Life

Technologies, a division of Invitrogen 9800 Medical Center

Drive Rockville, Maryland 20850, USA Fax: (1) 301 610

8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT

245 a 225 c 268 g 247 t 13 others

ORIGIN

```

QY 1 ValGluLeuArgGln-LeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAl 20
DB 414 GTAGACTTCGGAGAGTTATCCGCTGTGAACATCTTAATATTGTAAGCTTATGAGAC 473
QY 20 aCysLeuAsnProValCysLeuValMetGluTyr-AlaGluGlyLysSerLeuTyrAsnV 40
DB 474 CTGCTTGAATCCAGTGTGCTTGTGATGGAATATAGCTGAAGGGGCTCTTATATAATG 533
QY 40 alleuhsGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysL 60
DB 534 TGCTGATGCTGCTGAACCATTAATTAATTAATTAATTAATTAATTAATTAATTAATG 593
QY 60 euGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProGlyAlaLeuHisAs 80
DB 594 TACAGTTCCTCCAGAGAGGCTTATCTTCAAGATGCAACCAAGCTTATTCACA 653
QY 80 rGAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysA 100
DB 654 GGGACCTGGAACACCAACTTACTGCTGTGAGGGGGGACAGTCTTAAAAATTTGTG 713
QY 100 spPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaIaIaTr 120

```


Score: 823.00 Matches: 150
 Percent Similarity: 98.05% Conservative: 1
 Best Local Similarity: 97.40% Mismatches: 3
 Query Match: 65.73% Indels: 0
 DB: 14 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BQ386875 (1-641)

QY 75 LysAlaLeuIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThr 94
 Db 13 AAGGCTCTATTACAGAGACCTCAACCAACCACTTCCTGCTGGTACCTGGAGGACC 72
 QY 95 ValIleuLysIleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsn 114
 Db 73 GTTCTTAAGATTGTGACTTGTGACAGCTGGATATTCAAGCTCACATGACTATTAAC 132
 QY 115 LysGlySerAlaAlaTrpMetAlaProGluValPheGlyGlySerAsnTyrSerGluLys 134
 Db 133 AAGGAAAGTCAGACATGATGGCTCCAGAACTTTTGAAGGTAGCAATTACAGCAGAAA 192
 QY 135 CysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysPhe 154
 Db 193 TGTGACGATTATTAGTTGGGGAATATTCTTTGGAGATATAACCGAAGAAACCTTTC 252
 QY 155 AspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgPro 174
 Db 253 GATGAATAATGGTGTCCACGTTCCGTATATGTGGCTGTTCAACATGGTACTGCCCA 312
 QY 175 ProLeuIleLysAsnLeuProLysProIleGlySerLeuMetThrArgCysTrpSerLys 194
 Db 313 CCATTAAATTAATAAATTGGCTTAAGCTATTGAAGCTTAATGACTCGCTGCTGCAAA 372
 QY 195 AspProSerGlyArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArg 214
 Db 373 GATCCCCCAAAAGACCTTCAATGAGAGATTGTCAMATATATGATCATTAATGCA 432
 QY 215 TyrPheProGlyAlaAspGluProLeuGlnTyrProCysGln 228
 Db 433 TATTTTCTGGAGCGGATGTTCTTACAGTATCTTGTGCA 474

RESULT 11

AL550589 827 bp mRNA linear EST 16-FEB-2001
 LOCUS AL550589 LTI NFL006.PL2 Homo sapiens cDNA clone CS0D1058Y014 5
 DEFINITION prime, mRNA sequence.

ACCESSION AL550589
 VERSION AL550589.1 GI:12887700
 KEYWORDS EST.

SOURCE

ORGANISM human.

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)

COMMENT

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 Evry cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1..827
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="CS0D1058Y014"
 /clone_id="LTI NFL006.PL2"
 /tissue_type="Placenta"
 /note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA
 was primed with a NotI-oligo(dT) primer. Five prime end
 enriched, double-stranded cDNA was digested with Not I and
 cloned into the Not I and Eco RV sites of the pCMVSPORT 6
 vector. Library was normalized. Library was constructed by
 Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive
 Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
 Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/

BASE COUNT 196 a 178 c 228 g 209 t 16 others

ORIGIN

Alignment Scores:

Pred. No.: 1,85e-80 Length: 827
 Score: 787.00 Matches: 147
 Percent Similarity: 98.66% Conservative: 0
 Best Local Similarity: 98.66% Mismatches: 1
 Query Match: 62.86% Indels: 1
 DB: 9 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x AL550589 (1-827)

QY 1 ValGluLeuArgGluLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
 Db 382 GTAGAGCTTGGCAGATTATCCCGTGAACCATCTTAATTGTAAAGCTTTATGAGCC 441
 QY 21 CysLeuAsnProValCysLeuValMetGlyTyrAlaGlyGlySerLeuTyrAsnVal 40
 Db 442 TGCCTGAATCAGAGTGTCTTGTGATGGAATATGCTGAAGGGGCTCTTATATATGTR 501
 QY 41 LeuHisGlyAlaGluProLeuProTyrTyrTrpAlaAlaHisAlaMetSerTrpCysLeu 60
 Db 502 CTGCATGTGCTGAACCATTTGCATATTAATCTGCTGCCACGCAATGATGGTGTTTA 561
 QY 61 GlnCysSerGln-GlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
 Db 562 CAGGTTCACCAAGAGAGTGGCTTATCTTCACAGACATGCAACCAAGGCCCTTAATTCAC 621
 QY 80 GAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsn 100
 Db 622 GGAACCTGAAACCAACCAACTTACTGCTGTGAGGGGGAGCAGTTCTAATAATTGTGA 681
 QY 100 pPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTr 120
 Db 682 TTTTGTACAGCTGTGACATTTAGACACATGACCAATTAACAAGGGAGTGTGCTTGG 741
 QY 120 pMetAlaProGluValPheGluGlySerAsnTyrSerGlyLysCysAspValPheSerTr 140
 Db 742 GATGGCACCTCGAAGTTTGTGAAGTATTAATTAAGTGAATAAATGTGACGTTCACTG 801
 QY 140 pGlyIleIleLeuTyrGluValIle 148
 Db 802 GGGTATTATTCTTTGGGAAGTGATA 826

RESULT 12

BQ219348 1006 bp mRNA linear EST 02-MAY-2002
 LOCUS BQ219348
 DEFINITION AGENCOURT 7260864 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5785294
 5', mRNA sequence.

ACCESSION BQ219348
 VERSION BQ219348.1 GI:20400748
 KEYWORDS EST.

SOURCE

ORGANISM human.

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: c99abs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA library Preparation: Life Technologies, Inc.
 DNA library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:

http://image.llnl.gov
Plate: LLAMI2872 row: c column: 23
High quality sequence stop: 624.

FEATURES
source
1..1006
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5785294"
/clone_lib="NIH MGC 71"
/tissue type="leiomyosarcoma"
/lab host="DH10B (phage-resistant)"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr.
Average insert size 2.1 kb. "

BASE COUNT 240 a 240 c 293 g 233 t
ORIGIN

Alignment Scores:
Pred. No.: 2,16e-76 Length: 1006
Score: 753.00 Matches: 151
Percent Similarity: 78.39% Conservative: 5
Best Local Similarity: 75.88% Mismatches: 17
Query Match: 60.14% Indels: 27
DB: 14 Gaps: 4

US-09-830-144-2_COPY_76_303 (1-228) x BQ219348 (1-1006)

QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAla 20
DB 388 GTAGAGCTTCGGCAGTTATCCGCTGGAACATCCTTAATTTGAAGCTTTATGAGCC 447
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
DB 448 TGCTTGAATCCAGTGTCTCTTGATGAATATGCTGAAGGGGCGCTTTATATATATG 507
QY 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCysLeu 60
DB 508 CTGCATGTGTGAACCATTTGCCATATTATCTGCTGCCACGCAATAGTTGTTGTTTA 567
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
DB 568 CAGTGTTCACAGAGTGGCTTATCTTCAGCATGCAACCAAGCGCTAATTCACAGG 627
QY 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
DB 628 GACCTGAACACACCACTTACTGCTGTTGCGGGGGGACAGTTCT-AAAATTTGTGAT 686
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaAlaTrp 120
DB 687 TTGGGTACAGCTGTGACTTTCACACACATGACCAATTTACAAGGGGAGTGTGCTTGG 746
QY 121 MetAlaProGluValPheGluGlySerAsnTyrSerGlnLysCysAspValPheSerTrp 140
DB 747 ATGCACCTCGAATTTTGGAGGAGTAACCCGTCGGGAACCCCTTTGAGAGAGATTGGGG 806
QY 141 GlyIleIleLeuTrpGluValIleThrArgArg---LysProPheAsp---GluIleGly 158
DB 807 GGGAAATTTCTTTGGGAAGGATAACCCGTCGGGAACCCCTTTGAGAGAGATTGGGG 866
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys 178
DB 867 GGGCAACTTTT-----
QY 179 AsnLeuProLysProIle-----GluSerLeuMetThrArgCys 191
DB 879 -----CCAAACCACTGGGGCGGGGCGGCTCATACATGATGACCCCAACCCCGCTGC 929

RESULT 13
BG548917 602575750F1 NIH_MGC_77 Homo sapiens cdna clone IMAGE:4703580 5',
LOCUS mRNA sequence.
DEFINITION BG548917
ACCESSION

BG548917.1 GI:13547582
EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 686)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cdna Library Preparation: CLONTECH Laboratories, Inc.
cdna Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCMI542 row: d column: 13
High quality sequence stop: 671.

FEATURES
source
1..686
Location/Qualifiers

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4703580"
/clone_lib="NIH_MGC_77"
/lab host="DH10B (TI phage-resistant)"
/note="Organ: lung; Vector: pNR-LIB (Clontech); Site 1:
SfiI (ggcgccctcgcc); Site 2: SfiI (ggccattggcc); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGAGCGCGGCATG-dt(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH MGC Library."

BASE COUNT 202 a 143 c 169 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 4.5e-76 Length: 686
Score: 748.00 Matches: 144
Percent Similarity: 95.39% Conservative: 1
Best Local Similarity: 94.74% Mismatches: 4
Query Match: 59.74% Indels: 3
DB: 12 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BG548917 (1-686)

QY 80 ArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCys 99
DB 4 AGGGACCTGAAACCAACCAACTTACTGCTGTAGCTGGGGGACAGTTCTAAAAATTTGT 63
QY 100 AspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLysGlySerAlaAla 119
DB 64 GATTTTGTGTACAGCTGTGACATTCAGACACATGACCAATAACAGGGAGTGTGCT 123
QY 120 TrpMetAlaProGluValPheGluGlySerAsnTyrSerGlu-LysCysAspValPheSe 139
DB 124 TGGATGGCACCTGAAGTTTGTGAAGGTAGTAATTACAGTGAACAATGTCAGCTTCAG 183
QY 139 rTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyGl 159
DB 184 CTGGGGTATTATTCTTTGGGAAGTGATAACCGCTCGGAACCCCTTTGATGAGATTGGTGG 243
QY 159 yProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLys-A 179
DB 244 CCCACCTTTCCGAATCATGTGGGCTGTTCAATGTTGTTACTCGNACAGCACTGATAACAA 303
QY 179 snLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGln- 198

Db 304 ATTACTTAAGCCCTTGAAGCCTGATGACTGTGTTGCTCTAAGATCCTCCAGG 363

Qy 199 AagpGserMeGlUglUileValysIleMeThHisLeuMeArGtyrPheProGly 218

Db 364 CGCCCTTCATGAGGAAATTTGAAATAATGACTCACTTATGCGGACTTCCAGGA 423

Qy 219 AlaAspGluProLeuGlnTyPrCoGyGln 228

Db 424 GCAGATGAGCCATTACAGTATCCTGTGAG 453

RESULT 14

BI093821 771 bp mRNA linear EST 20-JUN-2001

LOCUS 602860601F1 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:5001504 5',

DEFINITION mRNA sequence.

ACCESSION BI093821

VERSION BI093821.1 GI:14512151

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 771)

TITLE NIH-MGC http://mgi.nci.nih.gov/.

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgaabs-r@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: Incyte Genomics, Inc.

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LLM11035 row: b column: 01

High quality sequence stop: 768.

FEATURES

source

1..771

Location/Qualifiers

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5001504"

/clone_1lb="NIH_MGC_10"

/cell_line="MGC36"

/lab_host="DH10B"

/note="Organ: cervix; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.5 kb. Library prepared by Life Technologies."

BASE COUNT 230 a 163 c 187 g 191 t

ORIGIN

Alignment Scores:

Pred. No.: 3,49e-75 Length: 771

Score: 741.00 Matches: 133

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 59.19% Indels: 0

DB: 13 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x BI093821 (1-771)

Qy 96 LeuLysIleCysAspPheGlyThralaCysAspIleGlnThHisMeThHisAsnLys 115

Db 1 CTAATAATTTGATTTTGGTACAGCCTGTGACATTCACACACATGACCAATAACAG 60

Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTySerGluLysCys 135

Db 61 GGAGAGTGCGCTTGATGACCACTGAAGTTTGAAGGTAGTAATTAACAGTAAATAATGT 120

Qy 136 AspValPheSerTrpGlyLleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155

Db 121 GACGCTTCAGTGCGGTATTAATTTCTTGGGAAGTATACCGGTGGAACCTTTGAT 180

Qy 156 GlUileGlyLysProAlaPheArgIleMeThrPAlaValHisAsnGlyThrArgProPro 175

Db 181 GAGATGTGTGGCCCAAGCTTTCGAATCATGTGGCTGTTTCATATGATGACTCCACCACA 240

Qy 176 LeuIleLysAsnLeuProLysProIleGluSerLeuMeThrArgCysTrpSerLysAsp 195

Db 241 CTGATTAATAAATTTACTTAAGCCATTGAGAGCCTGATGACTGTGTGTGTTAAAGAT 300

Qy 196 ProSerGlnArgProSerMeGlUglUileValysIleMeThHisLeuMeArGtyr 215

Db 301 CTTCCAGCGCCCTTCATGAGGAATTTGAAATAATGACTCACTGATGAGCGGTAC 360

Qy 216 PheProGlyAlaAspGluProLeuGlnTyPrCoGyGln 228

Db 361 TTTCAGAGAGAGATGAGCCATTACAGTATCCTGTGAG 399

RESULT 15

AV398933 682 bp mRNA linear EST 05-FEB-2000

LOCUS AV398933

DEFINITION AV398933 Bombyx mori ovary BmNPV infected; 6 hr after inoculation

ACCESSION AV398933

VERSION AV398933.1 GI:6902585

KEYWORDS EST.

SOURCE domestic silkworm.

ORGANISM Bombyx mori

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Dipteryia; Bombycoidea; Bombycidae; Bombyx.

AUTHORS 1 (bases 1 to 682)

TITLE Mita, K., Morimoto, M., Shimada, T., Okano, K. and Maeda, S.

JOURNAL Bombyx mori cDNA

COMMENT Unpublished (2000)

Contact: Mita K

Genome Research Group

National Institute of Radiological Sciences

Anagawa 4-9-1, Inage, Chiba 263-8555, Japan

Email: kmkita@nirs.go.jp

method:uni-directional, sequence direction:sequenced from T3 primer (5' -> 3')

Project="Silkworm Genome Program in MAF, and Research for the Future Program in USFS". see 'SilkBase', <http://www.ab.a.u-tokyo.ac.jp/silkbase/>, for whole ESTdb.

FEATURES

source

1..682

Location/Qualifiers

/organism="Bombyx mori"

/db_xref="taxon:7091"

/clone="NV060441"

/clone_1lb="Bombyx mori ovary BmNPV infected; 6 hr after inoculation"

/tissue_type="ovary"

/cell_type="BmNPV infected cell"

/dev_stage="BmNPV infected; 6 hr after inoculation"

BASE COUNT 200 a 130 c 191 g 160 t

ORIGIN

Alignment Scores:

Pred. No.: 8.37e-75 Length: 682

Score: 737.00 Matches: 139

Percent Similarity: 85.80% Conservative: 12

Best Local Similarity: 78.98% Mismatches: 21

Query Match: 58.87% Indels: 3

DB: 10 Gaps: 3

US-09-830-144-2_COPY_76_303 (1-228) x AV398933 (1-682)

Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrgIala 20

Db 159 ATCAGAGTACGCCAGTATATCTCGGTATGTCACCAAACTTGTGGATTTGACGAGCG 218

Qy 21 CysLeuAsn-----ProValCysLeuValMeGlUtyrAlaGluGlySerLeuTy 38


```
Db 219 TGCACGCGGGGTCTCACGTATGCCTCGTCATGGAGTATGCAGAAGGTGGCTCTCTCTAT 278
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrThrAlaAlaHisAlaMetSerTrp 58
Db 279 AATGTACTACCAAC---CGACCCAAACCCAAATACACAGCAGCACATGCAATGAGCTGG 335
QY 59 CysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIle 78
Db 336 GCCCGCCAGTGGCTGAGGGTGTAGCATATCTTCACTCCATGAACCAAAACCATTAATA 395
QY 79 HisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIle 98
Db 396 CACAGGGATTAAACCAACCAATCTATTATTAGTGGGTGGAGGCAACGGCTGAAGATT 455
QY 99 CysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAla 118
Db 456 TGTGATTTTGGCACAGCAGCTGATAAGGCAACATACATGACCAATAATAAAGGCAGTGCT 515
QY 119 AlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPhe 138
Db 516 GCTTGGATGGCTCTGAGGTATTGAAGGATCGACATACACAGAAAAAATGTGATGTGTC 575
QY 139 SerTrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAspGluIleGly 158
Db 576 TCATGGGGGATAATACTGTGGGAGGTATTATCTAGGAGGAAACCATTTGAAGAA---GGT 632
QY 159 GlyProAlaPheArgIleMetTrpAlaValHisAsnGlyThrArgPro 174
Db 633 GGTTCTGCATATAGAATAATGTGGCGGTGTCATACAGGTCNAAGACCA 680
```

Search completed: December 10, 2002, 05:34:37
Job time : 2223 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 10, 2002, 03:59:35 ; Search time 68 Seconds

(without alignments)
1028.269 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHPTIKLYGA.....MTHLMRYFPGADEPLQVPCQ 228

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapop 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141755_4367/app_query.fasta_1.391
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORW=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144@cgn_1_1_31@runat_04122002_141755_4367 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTCUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	1788	4	US-09-529-279-14 Sequence 14, Appli
2	1252	100.0	2856	2	US-08-685-625A-5 Sequence 5, Appli
3	1252	100.0	2856	4	US-09-529-279-3 Sequence 3, Appli
4	1247	99.6	2443	3	US-08-685-625A-1 Sequence 1, Appli
5	371	29.6	1365	3	US-09-221-235-6 Sequence 6, Appli
6	371	29.6	1365	3	US-09-221-928-6 Sequence 6, Appli
7	371	29.6	1365	3	US-09-221-527-6 Sequence 6, Appli
8	371	29.6	1365	3	US-09-221-236-6 Sequence 6, Appli
9	371	29.6	1365	3	US-09-221-416-6 Sequence 6, Appli
10	371	29.6	1365	4	US-09-221-245-6 Sequence 6, Appli
11	371	29.6	1365	4	US-09-163-115-6 Sequence 6, Appli
12	371	29.6	1365	4	US-09-221-528-6 Sequence 6, Appli

13	371	29.6	1365	4	US-09-593-553-6 Sequence 6, Appli
14	371	29.6	1365	4	US-09-221-237-6 Sequence 6, Appli
15	371	29.6	2120	3	US-09-221-235-4 Sequence 4, Appli
16	371	29.6	2120	3	US-09-221-928-4 Sequence 4, Appli
17	371	29.6	2120	3	US-09-221-527-4 Sequence 4, Appli
18	371	29.6	2120	3	US-09-221-236-4 Sequence 4, Appli
19	371	29.6	2120	3	US-09-221-416-4 Sequence 4, Appli
20	371	29.6	2120	4	US-09-221-245-4 Sequence 4, Appli
21	371	29.6	2120	4	US-09-163-115-4 Sequence 4, Appli
22	371	29.6	2120	4	US-09-221-528-4 Sequence 4, Appli
23	371	29.6	2120	4	US-09-593-553-4 Sequence 4, Appli
24	371	29.6	2120	4	US-09-221-237-4 Sequence 4, Appli
25	353.5	28.2	3389	1	US-08-395-580-1 Sequence 1, Appli
26	353.5	28.2	3426	1	US-08-205-018-1 Sequence 1, Appli
27	338.5	27.0	2505	4	US-09-291-839-3 Sequence 3, Appli
28	338.5	27.0	3025	4	US-09-291-839-1 Sequence 1, Appli
29	323	25.8	2890	1	US-07-928-464-1 Sequence 1, Appli
30	323	25.8	2890	5	PCT-US93-07347-1 Sequence 1, Appli
31	323	25.8	3033	1	US-08-003-311B-1 Sequence 1, Appli
32	323	25.8	3033	1	US-08-261-432-1 Sequence 1, Appli
33	305.5	24.4	3516	3	US-09-188-930-257 Sequence 257, Appl
34	304.5	24.3	4508	5	PCT-US93-06251-34 Sequence 34, Appl
35	294	23.5	1888	3	US-09-188-930-66 Sequence 66, Appl
36	293.5	23.4	2574	4	US-09-142-529-2 Sequence 2, Appli
37	280	22.4	1875	1	US-08-070-165F-3 Sequence 3, Appli
38	280	22.4	1875	2	US-08-885-418-3 Sequence 3, Appli
39	278	22.2	933	2	US-08-701-191A-4 Sequence 4, Appli
40	278	22.2	1056	2	US-08-701-191A-5 Sequence 5, Appli
41	278	22.2	2469	1	US-07-997-133-2 Sequence 2, Appli
42	278	22.2	2469	5	US-07-997-133-2 Sequence 2, Appli
43	278	22.2	2662	2	US-08-451-822A-14 Sequence 14, Appl
44	278	22.2	2662	4	US-08-323-430-14 Sequence 14, Appl
45	275	22.0	2469	1	US-08-459-296-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-529-279-14 Application US/09529279
; Sequence 14, Appli
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529, 279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)
US-09-529-279-14

Alignment Scores:

Pred. No.: 2,3e+140 Length: 1788
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-09-529-279-14 (1-1788)

```
QY 1 ValGluLeuArgGluLeuSerArgValaAsnHisProAsnIleValLysLeuTyrgIyala 20
Db 232 GTAGAGCTTCGGAGATTATCCCTGTGAACCACTTAATGTAAAGCTTATATGAGGC 291
QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40
Db 292 TGGTTGAATCAGTGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTG 351
QY 41 LeuHisGlyAlaGluProLeuProTyryTyraAlaAlaHisAlaMetSerTrpCysLeu 60
Db 352 CTCGATGAGTCTGAACCATGTGATTAATGCTGCTGCCAGCAATGAGTTGGTGTGTTA 411
QY 61 GluCysSerGluGlyValaAlaTyreLeuHisSerMetGlnProValaLeuLeuHisArg 80
Db 412 CAGTGTTCCTCCAAAGAGTGGTTATCTTCAAGATGCAACCCAAACGGTAAATTCACAG 471
QY 81 AspLeuLysProProAsnLeuLeuValaAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 472 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGAGACGTTTAAAAATTTGTGAT 531
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIaIaTrp 120
Db 532 TTGGTACAGCCTGTGACATTCAGACACATGACCAATTAACAGGGAGTGTGCTGTGG 591
QY 121 MetAlaProGluValaPheGluGlySerAsnTyrySerGluLysCysAspAlaPheSerTrp 140
Db 592 ATGGCACCCTGAAGTTTGAAGGTAGTAATTAACAGTAAATGAGACCTCTTCAGCTGG 651
QY 141 GlyIleIleLeuTyrgIyValaIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 652 GGTATATCTTTGGGAAGTGAATGATACGCTCGAAACCTTTGATGAGATTGTGGCCCA 711
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 712 GCTTCCGATCATGTGGGCTGTTCATATGTACTCGACCACTGATTAATAAATTTTA 771
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
Db 772 CCTAACCCATTGAGAGCCTGATGACTGCTGTGTCTTAAAGATCTTCCAGAGCCCT 831
QY 201 SerMetGluGluIleValLysIleMetThrHisLeuMetArgTyryPheProGlyAlaAsp 220
Db 832 TCATGAGAGAAATGTGAAATAATGACTCACTGATCGGTACTTTCAGAGACAAAT 891
QY 221 GluProLeuGluTyryProCysGln 228
Db 892 GAGCCATTACAGTATCTGTGAG 915

RESULT 2
; Sequence 5, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UNO, Naoto
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; TITLE OF INVENTION: TRANSDUCTION SYSTEM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHEIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

```
APPLICATION NUMBER: US/08/685,625A
FILING DATE: 24-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 7-253549
FILING DATE: 29-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: Meuth, Donna M.
REGISTRATION NUMBER: 36,607
REFERENCE/DOCKET NUMBER: 001560-267
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2656 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 183..1922
US-08-685-625A-5

Alignment Scores:
Pred. No.: 4,17e-140 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 2

US-09-830-144-2_COPY_76_303 (1-228) x US-08-685-625A-5 (1-2656)
QY 1 ValGluLeuArgGluLeuSerArgValaAsnHisProAsnIleValLysLeuTyrgIyala 20
Db 408 GTAGAGCTTCGGAGATTATCCCTGTGAACCACTTAATGTAAAGCTTATATGAGGC 467
QY 21 CysLeuAsnProValCysLeuValMetGluTyraIaGluGlySerLeuTyraAsnVal 40
Db 468 TGGTTGAATCAGTGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATATGTG 527
QY 41 LeuHisGlyAlaGluProLeuProTyryTyraAlaAlaHisAlaMetSerTrpCysLeu 60
Db 528 CTGATGTGTCTGAACCATTCATTAATTAATGCTGCCACCAATGAGTGTGTTTA 587
QY 61 GluCysSerGluGlyValaAlaTyreLeuHisSerMetGlnProLysAlaLeuLeuHisArg 80
Db 588 CAGTGTTCCTCCAAAGAGTGGTTATCTTCAAGCATGCAACCCAAAGCGCTTAATTCACAG 647
QY 81 AspLeuLysProProAsnLeuLeuValaAlaGlyGlyThrValLeuLysIleCysAsp 100
Db 648 GACCTGAACCAACCAACTTACTGCTGTGTCAGGGGGAGACGTTTAAAAATTTGTGAT 707
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaIaIaTrp 120
Db 708 TTGGTACAGCCTGTGATCAATTCAGACACATGACCAATTAACAAAGGGAGTGTGCTGG 767
QY 121 MetAlaProGluValaPheGluGlySerAsnTyrySerGluLysCysAspAlaPheSerTrp 140
Db 768 ATGGCACCCTGAAGTTTGAAGGTAGTAATTAACAGTAAATGTAACGCTTTCAGCTGG 827
QY 141 GlyIleIleLeuTyrgIyValaIleThrArgArgLysProPheAspGluIleGlyGlyPro 160
Db 828 GGTATATATTTTGGGAAGTGAATGAGCGTCGAAACCTTTGATGAGATTGTGGCCCA 887
QY 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
Db 888 GCTTCCGATCATGTGGGCTGTTCATATATGATGACTCGACCACTGATTAATAAATTTTA 947
QY 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
```

Db 948 CCTAAGCCATTGAGAGCCTGATGACTCGTTGGTGTCTAAAGATCCTCCCGAGCCCT 1007
Qy 201 SerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
|||||
Db 1008 TCAATGGAGGAATTGTGAAATTAATGACTCATTGTATGGGTACTTTCCAGGAGCAT 1067
Qy 221 GluProLeuGlnTyrProCysGln 228
|||||
Db 1068 GAGCCATTACAGTATCCTTGTTCAG 1091
RESULT 3
US-09-529-279-3
; Sequence 3, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-09-529-279-3
Alignment Scores:
Pred. No.: 4.17e-140 Length: 2656
Score: 1252.00 Matches: 228
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 4
US-09-830-144-2_COPY_76_303 (1-228) x US-09-529-279-3 (1-2656)
Qy 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIleLysLeuTyrGlyAla 20
|||||
Db 408 GTAGAGCTTCGGCAGTTATCCCGTGTGAACCATCCTTAATATTGTAAGCTTTATGGAGCC 467
Qy 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsnVal 40
|||||
Db 468 TGCTTGAATCCAGTGTCTTGTGATGGAATATGCTGAAGGGGGCTCTTTATATATATG 527
Qy 41 LeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
|||||
Db 528 CTGCATGGTGTCTGAACCATTTGCCATATTATCTGCTGCCACCAATGATGTTGTTTA 587
Qy 61 GlnCysSerGlnGlyValIleTyrLeuHisSerMetGlnProLysAlaLeuIleHisArg 80
|||||
Db 588 CAGTGTTCACAGGAGTGGCTTATCTTCAGCATGCAACCCAAAGCGCTAATTCACAGG 647
Qy 81 AspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLysIleCysAsp 100
|||||
Db 648 GACCTGAACACCAAACTTACTGTGTTGTCAGGGGGGACAGTTCTAAAAATTTGTGAT 707
Qy 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
|||||
Db 708 TTTGGTACAGCCTGTGACATTCAGACACACATGACCAATCAAGGGGAGTCTGCTTGG 767
Qy 121 MetAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAspValPheSerTrp 140
|||||

Db 768 ATGGCACCTGAAGTATTTTGAAGTAGTAATTTACAGTGAATAATGTGACGTCTTCAGCTGG 827
Qy 141 GlyIleIleLeuTrpGluValIleThrArgArgLysProPheAspGluIleGlyClyPro 160
|||||
Db 828 GGTATTATTCTTTGGGAAGTGATAACGCGTCGGAACCCTTTGTATGAGATTGGTGGCCCA 887
Qy 161 AlaPheArgIleMetTrpAlaValHisAsnGlyThrArgProProLeuIleLysAsnLeu 180
|||||
Db 888 GCTTTCCGAATCATGTGGCTGTTCAATAATGTACTGACCACCACCTGATAAAAATTTA 947
Qy 181 ProLysProIleGluSerLeuMetThrArgCysTrpSerLysAspProSerGlnArgPro 200
|||||
Db 948 CCTAAGCCATTGAGAGCCTGATGACTCGTTGGTGTCTAAAGATCCTTCCAGCGCCCT 1007
Qy 201 SerMetGluGluIleValIleMetThrHisLeuMetArgTyrPheProGlyAlaAsp 220
|||||
Db 1008 TCAATGGAGGAATTGTGAAATTAATGACTCATTGTATGGGTACTTTCCAGGAGCAT 1067
Qy 221 GluProLeuGlnTyrProCysGln 228
|||||
Db 1068 GAGCCATTACAGTATCCTTGTTCAG 1091
RESULT 4
US-08-685-625A-1
; Sequence 1, Application US/08685625A
; Patent No. 5945301
; GENERAL INFORMATION:
; APPLICANT: UENO, Naoto
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: IRIE, Kenji
; TITLE OF INVENTION: NOVEL KINASE IN TGF-BETA FAMILY SIGNAL
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, L.L.P.
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/685,625A
; FILING DATE: 24-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 7-253549
; FILING DATE: 29-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Meuth, Donna M.
; REGISTRATION NUMBER: 36,607
; REFERENCE/DOCKET NUMBER: 001560-267
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2443 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 157..1893
US-08-685-625A-1
Alignment Scores:
Pred. No.: 1.46e-139 Length: 2443

Score: 1247.00 Matches: 227
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 99.60% Indels: 0
DB: 2 Gaps: 0

US-09-830-144-2_COPY_76_303 (1-228) x US-08-685-625A-1 (1-2443)

```
QY 1 ValGluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAla 20
DB 382 GTGGAGACTCCGGCAGTGTGCGCGTGAACCATCTTAACATTGTCAAGTTGACGAGCC 441
QY 21 CysLeuAsnProValCysLeuValMetGluTyrAlaGlnGlyGlySerLeuTyrAsnVal 40
DB 442 TGCCTAAATCCAGTATGCTCTTGATGGAATATGCAGAGGGGGCTCATTTGATATGTTG 501
QY 41 LeuHisGlyValGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTyrCysLeu 60
DB 502 CTGCATGCTGCTGACCACTTGCTCTTACTACACTGCTGCTCATGCCATGAGCTGGTGTTA 561
QY 61 GlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeuIleHisArg 80
DB 562 CACTGTTCCCAAGAGAGTGGCTTACTGCACAGCATGCACCCCAAGCGCTGATTACAGG 621
QY 81 AspleuIysProProAsnLeuLeuLeuValAlaGlyGlyThrValLeuIysIleCysAsp 100
DB 622 GACCTCAAGCTTCCAACTTGCTGCTGCTGGTGCAGAGGAGCAGTTCTAAAAATCTGCCAT 681
QY 101 PheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIysGlySerAlaAlaTyr 120
DB 682 TTTGGTACACTTGTGACATCCAAACACACATGACCAATATTAAGGAGTGTCTGCTTG 741
QY 121 MetAlaProGluValPheGlnGlySerAsnTyrSerGlnIysCysAspValPheSerTyr 140
DB 742 ATGGCGCGCTGAAGTGTGTAAGGTAGCAATTACATGAAGAGTGTATCTTCACTGAG 801
QY 141 GlyIleIleLeuTyrGluValIleThrArgArgIysProPheAspIuIleGlyGlyPro 160
DB 802 GGTATATCTCTCTGGAGAGATGATACACGCCGGAACCCCTTCATAGATCGGTGCCCA 861
QY 161 AlaPheArgIleMetTyrPalaValHisAsnGlyThrArgProProLeuIleIysAsnLeu 180
DB 862 GCTTTTCAGATTCATGTGGCGCTGTTCAATATGACACTGCACCACTGATCAAAATTTA 921
QY 181 ProIysProIleGluSerLeuMetThrArgCysTyrSerIysAspProSerGlnArgPro 200
DB 922 CTTAAGCCCATTTGAGAGCTTGATGACACGCGTGTGCTTAAGAACCATCTCAGCGCCCT 981
QY 201 SerMetGluGluIleValIysIleMetThrHisLeuMetArgTyrPheProGluAlaAsp 220
DB 982 TCAATGAGAGAAATTGTGAAATTAATGACCTTGAATGCGTACTTCCAGAGCGGAT 1041
QY 221 GluProLeuGlnTyrProCys 227
DB 1042 GAGCCATTACATGATCTTGT 1062
```

RESULT 5
US-09-221-235-6
; Sequence 6, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA

ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1..(1365)
US-09-221-235-6

Alignment Scores:

Pred. No.:	6,566-35	Length:	1365
Score:	371.00	Matches:	81
Percent Similarity:	56.128%	Conservative:	40
Best Local Similarity:	37.67%	Mismatches:	80
Query Match:	29.63%	Indels:	14
DB:	3	Gaps:	8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-235-6 (1-1365)

```
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
DB 157 GAGGACAGAAATACTCACTGCTCCTCAGTACAGAAACATCATCCAGTTTATGAGATTAAT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGlnGlyGlySerLeuTyrAsn 39
DB 217 CTTGAACTCCCAACTATGAGCATTTGCACAAATATGCTTCTGTGGATCCTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
DB 277 TACATTAACTAACACAGAGTGAAGATG-----GATATGATCATTATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProIysAlaLeu 77
DB 328 TGGGCACTGATGTAGCCAAAGGAATGATTTATTAATGAGAGCTCTGCAAGGTG 387
QY 78 IleHisArgAspleuIysProProAsnLeuLeuValAlaGlyGlyThrValLeuIys 97
DB 388 ATTCAACAGAGACCTCAAGTCAAGAAAGTGTATTAAGTCTGATGGA-----GTAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnIys----- 115
DB 445 ATCTGATCTTTGAT---GCCCTCGGTTCCATTAACATACAAACACATGCTCTTGTT 501
QY 116 GlySerAlaAlaTyrMetAlaProGluValPheGlnGlySerAsnTyrSerGlnIysCys 135
DB 502 GGAACCTTCCCATGAGTGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561
QY 136 AspValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgIysProPheAsp 155
DB 562 GACACATATTCCTATGTGGTGGTCTCTGGAGATGCTTAACAAGGAGGTCCCTTTAAA 621
QY 156 GluIleGlyGlyProAlaPheArgIleMetTyr--AlaValHisAsnGlyThrArgPro 174
DB 622 GGTITGAAAGGA-----TTACAAAGTACGTTGGCTTATGTGAAAAAAGAGATTA 675
QY 175 ProLeuIleIysAsnLeuProIysProIleGluSerLeuMetThrArgCysTyrSerIys 194
DB 676 ACAAATCCAGCAGTGTGGCCAGAAAGTTTGTGTAAGTGTATCACTGATGTGGAGAGT 725
QY 195 AspProSerGlnArgProSerMetGluGlnIleValIysIleMet 209
DB 736 GATGCCAAGAAACGGCCATCATTCACAAATCAATTCATCCG 780
```

RESULT 6
US-09-221-928-6
; Sequence 6, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1365)

US-09-221-928-6

Alignment Scores:

Pred. No.: 6.56e-35 Length: 1365

Score: 371.00 Matches: 81

Percent Similarity: 56.28% Conservative: 40

Best Local Similarity: 37.67% Mismatches: 80

Query Match: 29.63% Indels: 14

DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-928-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21

Db 157 GAGCAGAAATCTACTAGTCTCAGTCACAGAAACATCCAGTTTATGGAGTAATT 216

QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39

Db 217 CTTGAACCTCCCAACTATGGCATTGTCACAGAAATATGCTTCTCTGGGATCACTCTATGAT 276

QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57

Db 277 TACATTAACTACAGAAAGTGAGGAGATG-----GATATGGATCAATATATGACC 327

QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77

Db 328 TGGGCCACTGATGTAGCCAAAGGAATGCATTATTTACATATGAGGCTCTCTCAAGGTG 387

QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97

Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAG 444

QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115

Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACAACACACATGCTCCTGGTT 501

QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlyLysCys 135

Db 502 GGAACCTTCCCATGGATGGCTCCAGAGATTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561

QY 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgGlySerProPheAsp 155

Db 562 GACACATATTCCTATGCTGTGTTCTCTGGAGATGCTAAACAGGGAGGTCCTTTAAA 621

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174

Db 622 GGTTTGGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACAAGAGATTA 675

QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194

Db 676 ACCATTCCAAGCAGTTCGCCAGAGATTTTGTGGAACGTGTTACATCAGTGTGGGAAGCT 735

QY 195 AspProSerGlnArgProSerMetGluIleValLysIleMet 209

Db 736 GATGCCAAGAAACGCCCATCATTCGAAGAAATCATTTCAATCCTG 780

RESULT 7

US-09-221-527-6

; Sequence 6, Application US/09221527

; Patent No. 6146832

; GENERAL INFORMATION:

; APPLICANT: Acton, Susan

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

; FILE REFERENCE: MNI-050

; CURRENT APPLICATION NUMBER: US/09/221,527

; CURRENT FILING DATE: 1998-12-28

; EARLIER APPLICATION NUMBER: 09/163,115

; EARLIER FILING DATE:

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1365

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1365)

US-09-221-527-6

Alignment Scores:

Pred. No.: 6.56e-35 Length: 1365

Score: 371.00 Matches: 81

Percent Similarity: 56.28% Conservative: 40

Best Local Similarity: 37.67% Mismatches: 80

Query Match: 29.63% Indels: 14

DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-527-6 (1-1365)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21

Db 157 GAGCAGAAATCTACTAGTCTCAGTCACAGAAACATCCAGTTTATGGAGTAATT 216

QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39

Db 217 CTTGAACCTCCCAACTATGGCATTGTCACAGAAATATGCTTCTCTGGGATCACTCTATGAT 276

QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57

Db 277 TACATTAACTACAGAAAGTGAGGAGATG-----GATATGGATCAATATATGACC 327

QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77

Db 328 TGGGCCACTGATGTAGCCAAAGGAATGCATTATTTACATATGAGGCTCTCTCAAGGTG 387

QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97

Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTTATAGCTGCTGATGGA---GTACTGAAG 444

QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115

Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCATACAACACACATGCTCCTGGTT 501

QY 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGlyLysCys 135

Db 502 GGAACCTTCCCATGGATGGCTCCAGAGATTATCCAGAGTCTCCCTGTGTCAGAAACTTGT 561

QY 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgGlySerProPheAsp 155

Db 562 GACACATATTCCTATGCTGTGTTCTCTGGAGATGCTAAACAGGGAGGTCCTTTAAA 621

QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174

Db 622 GGTTTGGAAGGA-----TTACAAGTAGCTTGGCTTGTAGTGAAAAAACAAGAGATTA 675

QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194

Db 676 ACCATTCCAAGCAGTTCGCCAGAGATTTTGTGGAACGTGTTACATCAGTGTGGGAAGCT 735

QY 195 AspProSerGlnArgProSerMetGluIleValLysIleMet 209

Db 736 GATGCCAAGAAACGCCCATCATTCGAAGAAATCATTTCAATCCTG 780

RESULT 8

US-09-221-236-6

; Sequence 6, Application US/09221236

; Patent No. 6146841

; GENERAL INFORMATION:

```

; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; EARLIER FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-236-6

```

```

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: Gaps: 8

```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-236-6 (1-1365)

```

QY 2 G|L|e|u|a|r|g|l|e|u|s|e|r|a|r|g|v|a|l|a|s|h|i|p|r|o|a|s|h|i|e|v|a|l|y|s|e|u|r|y|g|l|a|c|y|s 21
Db 157 GAGGCGAAGTAATCTCACTGCTCCAGTCACAGAAACATCTCACTGATTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
Db 217 CTGGAACCTCCCAACTAGTGCATTTGCACAGAAATATGCTTCTCGGATCATCTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db 277 TACATTAACTGTAACAGAAAGTGAAGATG-----GATATGATCATATTATGACC 327
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGGCCACTGATGATGACCAAGAAATGCAATTATTACATATGAGGCTCCTGCAAGGTG 387
QY 78 I|L|E|H|S|A|R|G|A|P|L|E|U|Y|S|P|R|O|A|S|N|L|E|U|L|E|U|V|A|L|A|G|L|Y|T|H|R|V|A|L|L|E|U|Y|S 97
Db 388 ATTACAGAGAACCTCAAGTCAAGAAACGTTTATAGCTGATGGA---GTAAGTGAAG 444
QY 98 I|L|E|C|Y|S|A|P|H|E|G|L|Y|T|H|R|A|L|C|Y|S|A|S|P|L|E|G|I|N|T|H|I|S|E|T|H|R|A|S|N|L|Y|S----- 115
Db 445 ATCTGTGACTTTGGT---GCCTCGGGTTCATTAACCATCAACACACATGTCCTTGTT 501
QY 116 G|L|S|e|r|A|l|a|T|r|m|e|l|a|P|r|o|G|l|u|V|a|l|P|h|e|G|l|U|G|L|Y|S|e|r|a|n|T|y|r|S|e|r|G|l|u|Y|S|C|Y|S 135
Db 502 GGAACCTTCCCATGAGAGGCTCCAGAAATATCCAGAGTCTCCCTGTGTCAGAACTTGT 561
QY 136 A|s|p|V|a|l|P|h|e|S|e|r|T|r|p|G|l|Y|L|e|l|e|u|r|T|r|p|G|l|U|V|a|l|I|e|H|R|A|R|G|A|G|L|Y|S|P|R|O|P|h|e|A|S|P 155
Db 562 GACACATATTCCTATGTGTGTCTCTGGAGATGCTAACAGAGAGGCTCCCTTTAA 621
QY 156 G|L|U|I|e|G|L|Y|P|R|O|A|L|P|h|e|A|R|G|L|E|T|T|r|p|---A|L|a|V|A|H|I|S|A|N|G|L|Y|T|H|R|A|R|G|P|R|O 174
Db 622 GGTTTGGAAGA-----TTACAAGTACCTTGCTGTGTAAGTGAAGAAAAAAGAGAGATTGA 675
QY 175 P|r|o|L|e|u|I|e|L|y|S|A|S|N|L|e|U|P|r|o|L|y|S|P|r|o|I|e|G|L|S|e|r|L|e|U|e|T|H|R|A|R|G|C|Y|S|T|r|p|S|e|r|L|Y|S 194
Db 676 ACCATTCCAAAGAGTGGCCCAAGAAATTGTCGAACGTGTACATCAATGTTGGAGACT 735
QY 195 A|s|p|P|r|o|S|e|r|G|l|A|r|g|P|r|o|S|e|r|M|e|t|G|l|U|I|L|e|V|a|l|Y|S|I|L|E|T 209
Db 736 GATGCCAAGAAAGGCGCATCTTCAACGAATCATTTCAATCTCTG 780
RESULT 9

```

```

US-09-221-416-6
; Sequence 6, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-416-6

```

```

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: Gaps: 8

```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-416-6 (1-1365)

```

QY 2 G|L|e|u|a|r|g|l|e|u|s|e|r|a|r|g|v|a|l|a|s|h|i|p|r|o|a|s|h|i|e|v|a|l|y|s|e|u|r|y|g|l|a|c|y|s 21
Db 157 GAGGCGAAGTAATCTCACTGCTCCAGTCACAGAAACATCTCACTGATTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
Db 217 CTGGAACCTCCCAACTAGTGCATTTGCACAGAAATATGCTTCTCGGATCATCTATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db 277 TACATTAACTGTAACAGAAAGTGAAGATG-----GATATGATCATATTATGACC 327
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGGCCACTGATGATGACCAAGAAATGCAATTATTACATATGAGGCTCCTGCAAGGTG 387
QY 78 I|L|E|H|S|A|R|G|A|P|L|E|U|Y|S|P|R|O|A|S|N|L|E|U|L|E|U|V|A|L|A|G|L|Y|T|H|R|V|A|L|L|E|U|Y|S 97
Db 388 ATTACAGAGAACCTCAAGTCAAGAAACGTTTATAGCTGATGGA---GTAAGTGAAG 444
QY 98 I|L|E|C|Y|S|A|P|H|E|G|L|Y|T|H|R|A|L|C|Y|S|A|S|P|L|E|G|I|N|T|H|I|S|E|T|H|R|A|S|N|L|Y|S----- 115
Db 445 ATCTGTGACTTTGGT---GCCTCGGGTTCATTAACCATCAACACACATGTCCTTGTT 501
QY 116 G|L|S|e|r|A|l|a|T|r|m|e|l|a|P|r|o|G|l|u|V|a|l|P|h|e|G|l|U|G|L|Y|S|e|r|a|n|T|y|r|S|e|r|G|l|u|Y|S|C|Y|S 135
Db 502 GGAACCTTCCCATGAGAGGCTCCAGAAATATCCAGAGTCTCCCTGTGTCAGAACTTGT 561
QY 136 A|s|p|V|a|l|P|h|e|S|e|r|T|r|p|G|l|Y|L|e|l|e|u|r|T|r|p|G|l|U|V|a|l|I|e|H|R|A|R|G|A|G|L|Y|S|P|R|O|P|h|e|A|S|P 155
Db 562 GACACATATTCCTATGTGTGTCTCTGGAGATGCTAACAGAGAGGCTCCCTTTAA 621
QY 156 G|L|U|I|e|G|L|Y|P|R|O|A|L|P|h|e|A|R|G|L|E|T|T|r|p|---A|L|a|V|A|H|I|S|A|N|G|L|Y|T|H|R|A|R|G|P|R|O 174
Db 622 GGTTTGGAAGA-----TTACAAGTACCTTGCTGTGTAAGTGAAGAAAAAAGAGAGATTGA 675
QY 175 P|r|o|L|e|u|I|e|L|y|S|A|S|N|L|e|U|P|r|o|L|y|S|P|r|o|I|e|G|L|S|e|r|L|e|U|e|T|H|R|A|R|G|C|Y|S|T|r|p|S|e|r|L|Y|S 194
Db 676 ACCATTCCAAAGAGTGGCCCAAGAAATTGTCGAACGTGTACATCAATGTTGGAGACT 735
QY 195 A|s|p|P|r|o|S|e|r|G|l|A|r|g|P|r|o|S|e|r|M|e|t|G|l|U|I|L|e|V|a|l|Y|S|I|L|E|T 209

```


Db 736 GATGCCAAGAAACGGCCATCATTTCAAGCAATCATTTCAATCCTG 780

RESULT 10

```

US-09-221-245-6
; Sequence 6, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-245-6

```

Alignment Scores:	6.56e-35	Length:	1365
Pred. No.:	371.00	Matches:	81
Score:	56.28%	Conservative:	40
Percent Similarity:	37.67%	Mismatches:	80
Best Local Similarity:	29.63%	Indels:	14
Query Match:	4	Gaps:	8
DB:			

US-09-830-144-2 COPY 76 303 (1-228) x US-09-221-245-6 (1-1365)

QY	2	Glul	Arg	Gln	Ieu	Ser	Arg	Val	Asn	His	Pro	Asn	Ile	Val	Lys	Ieu	Tyr	Gly	Ala	Cys	21
DB	157	GAG	G	C	A	A	A	A	T	A	C	T	C	A	G	T	T	T	A	T	216
QY	22	Leu	Asn	Pro	-----	Val	Cys	Leu	Val	Met	Glu	Tyr	Ala	Alu	Gly	Ser	Ieu	Tyr	Asn	39	
DB	217	CTT	G	A	A	C	T	C	C	C	A	A	C	T	T	G	G	A	T	276	
QY	40	Val	Ile	His	Gly	-----	Ala	Glu	Pro	Leu	Pro	Tyr	Tyr	Thr	Ala	Ala	His	Ala	Met	57	
DB	277	TAC	A	T	T	A	A	C	A	G	T	A	G	G	A	G	A	T	G	327	
QY	58	Trp	Cys	Leu	Gln	Cys	Ser	Gln	Gly	Val	Ala	Tyr	Leu	His	Ser	Met	Gln	Pro	Lys	77	
DB	328	TGG	G	C	C	A	C	T	G	A	T	G	A	G	A	A	T	T	T	387	
QY	78	Ile	His	Arg	Asp	Ieu	Lys	Pro	Pro	Asn	Leu	Leu	Val	Ala	Gly	Thr	Val	Leu	Lys	97	
DB	388	ATT	C	A	G	A	G	A	C	T	C	A	A	G	T	T	T	A	T	444	
QY	98	Ile	Cys	Asp	Phe	Gly	Thr	Ala	Cys	Asp	Ile	Gln	Thr	His	Met	Thr	Asn	Asn	Lys	115	
DB	445	ATC	T	G	T	G	A	C	T	T	G	G	T	T	C	A	A	C	A	501	
QY	116	Gly	Ser	Ala	Ala	Trp	Met	Ala	Pro	Glu	Val	Phe	Glu	Gly	Ser	Asn	Tyr	Ser	Glu	135	
DB	502	GGA	A	C	T	T	C	C	A	T	G	G	T	T	A	T	C	C	A	561	
QY	136	Asp	Val	Phe	Ser	Trp	Gly	Ile	Ile	Leu	Trp	Glu	Val	Leu	Thr	Arg	Arg	Lys	Pro	155	
DB	562	GAC	A	C	A	T	A	T	T	C	T	A	T	G	T	G	G	T	T	621	
QY	156	Glu	Ile	Gly	Gly	Pro	Ala	Phe	Arg	Ile	Met	Trp	-----	Ala	Val	His	Asn	Gly	Thr	174	
DB	622	GGT	T	T	G	A	A	G	A	-----	T	T	A	A	G	T	A	G	T	675	
QY	175	Pro	Leu	Ile	Lys	Asn	Leu	Pro	Lys	Pro	Ile	Glu	Ser	Leu	Met	Thr	Arg	Cys	Trp	194	

Db 676 ACCATTCGAAGCAGTTGCCCGAAGAAGTTTGTCTGAACATGTATCATCTGTATGGGAAGCT 735

Qy 195 AspProSerGlnArgProSerMetGluIleValLysileMet 209
 ||| ::::| | | | | | : : : : : : : :
Db 736 GATGCCAAGAAACGGGCATCATTCAGAACAATCATTCAATCTGT 780

RESULT 11

```

US-09-163-115-6
; Sequence 6, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: WNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-163-115-6

```

Alignment Scores:		
Pred. No.:	6.56e-35	1365
Score:	371.00	81
Percent Similarity:	56.28%	Conservative: 40
Best Local Similarity:	37.67%	Mismatches: 80
Query Match:	29.63%	Indels: 14
DB:	4	Gaps: 8

US-09-830-144-2 COPY 76 303 (1-228) X US-09-163-115-6 (1-1365)

Qy	2	GlueuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrrGlyrAlaCys	21
Db	157	GAGGCAGAAATACTCAGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT	216
Qy	22	LeuAsnPro-----ValCysLeuValMetGluTyrrAlaGluGlySerLeuTyrrAsn	39
Db	217	CTTGAACCTCCAACTATGCATTTGTCACAGAAATATCTTCTCTGGGATCACTCTATGAT	276
Qy	40	ValLeuHisGly-----AlaGluProLeuProTyrrThrAlaAlaHisAlaMetSer	57
Db	277	TACATTTAACAGTACACAGAGTGGAGAGT-----GATATGGATCATATTATGACC	327
Qy	58	TrpCysLeuGlnCysSerGlnGlyValAlaTyrrLeuHisSerMetGlnProLysAlaLeu	77
Db	328	TGGGCCACTGATGTAGCCAAAGGATGCATATTATTACATATGAGGCTCTCTGTCAAGGTG	387
Qy	78	IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys	97
Db	388	ATTACACAGACCTCAAGTCAAGAAAGCTTTGTTATAGCTGCTGATGGA---GTACTGAAG	444
Qy	98	IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys-----	115
Db	445	ATCTGTGCATTTGGT---GGCTCTCGGTTCATTAACCATACACACATGTCCTTGGTT	501
Qy	116	GlySerAlaAlaTrpMetAlaProGluValPheGluCysSerAsnTyrrSerGluLysCys	135
Db	502	GGAACTTTCCCATGGATGGCTTCAGAAAGTTATTCAGAGTCTCCCTGTGTGAGAAACTTGT	561
Qy	136	AspValPheSerTrpGlyIleLeuLeuTrpGluValIleThrArgArgLysProPheAsp	155
Db	562	GACACATATTCTATGGTGTGGTTCTCTGGAGATGCTACACAGGGAGGTCCCTTTAA	621
Qy	156	GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro	174
Db	622	GGTTTGAAGA-----TTACAGTAGCTTGGCTTCTAGTGAAAAAAACGAGAGATTA	675

```
QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
DB 676 ACCATTCCAGACAGTGGCCCCAGAGTTTGGTGAACCTGTACATCAGTGGGAGACT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
DB 736 GATGCCAAGAAAGGCCCATCATTCAGCAAAATCATTTCAATCTCTG 780

RESULT 12
US-09-221-528-6
; Sequence 6, Application US/09221528
; Patent No. 6130874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-528-6

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 4 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-528-6 (1-1365)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrglyAlaCys 21
DB 157 GAGGCAGAAATATCTCAGTGTCTCCAGCTCAGCAAAATCATCATGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrglyAlaGluGlyGlySerLeuTyraen 39
DB 217 CTGGAACCTCCCAACTATGAGCATTTGTCAGAGAAATATGCTTCTCTGGGATCAGCTTAATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrglyThrAlaHisAlaMetSer 57
DB 277 TACATTAAACAGTAACAGAGTAGAGAGATG-----GATATGAGTACATTATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrglyLeuHisSerMetGlnProLysAlaLeu 77
DB 328 TGGGCACACTGATGTAGCCAAAGAGATGCAATTATTACATATGAGGCTCTCTGCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
DB 388 ATTCACAGAGACCTCAAGTCCAGAAAGCTGTATATGCTGCTGAGTGA--GTACTGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115
DB 445 ATCTGTGACTTTGGT---GCTCTCGGTTCCATTAACCATTAACAACACATGCTTTGGTT 501
QY 116 GlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrglyLysCys 135
DB 502 GGAACCTTCCCATGATGGCTCCAGAAATTATTCAGAGTCCCTGTCACAAAACCTTGT 561
QY 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
DB 562 GACACATATTCATTGATGTGTGTTCTCTGGAGATGCTAAACAAGGAGTCCCTTTTAA 621
```

```
QY 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
DB 622 GGTTTGGAAGA-----TTACAAGTAGCTTGCTGTGATGGAAAAAACAAGAGATTAA 675
QY 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
DB 676 ACCATTCCAGACAGTGGCCCCAGAGTTTGGTGAACCTGTACATCAGTGGGAGACT 735
QY 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
DB 736 GATGCCAAGAAAGGCCCATCATTCAGCAAAATCATTTCAATCTCTG 780

RESULT 13
US-09-593-553-6
; Sequence 6, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-593-553-6

Alignment Scores:
Pred. No.: 6,56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 4 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-593-553-6 (1-1365)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrglyAlaCys 21
DB 157 GAGGCAGAAATATCTCAGTGTCTCCAGCTCAGCAAAATCATCATGTTTATGAGTAATT 216
QY 22 LeuAsnPro-----ValCysLeuValMetGluTyrglyAlaGluGlyGlySerLeuTyraen 39
DB 217 CTGGAACCTCCCAACTATGAGCATTTGTCAGAGAAATATGCTTCTCTGGGATCAGCTTAATGAT 276
QY 40 ValLeuHisGly-----AlaGluProLeuProTyrglyThrAlaHisAlaMetSer 57
DB 277 TACATTAAACAGTAACAGAGTAGAGAGATG-----GATATGAGTACATTATGACC 327
QY 58 TyrCysLeuGlnCysSerGlnGlyValAlaTyrglyLeuHisSerMetGlnProLysAlaLeu 77
DB 328 TGGGCACACTGATGTAGCCAAAGAGATGCAATTATTACATATGAGGCTCTCTGCAAGGTG 387
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
DB 388 ATTCACAGAGACCTCAAGTCCAGAAAGCTGTATATGCTGCTGAGTGA--GTACTGAAG 444
QY 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnLys----- 115
DB 445 ATCTGTGACTTTGGT---GCTCTCGGTTCCATTAACCATTAACAACACATGCTTTGGTT 501
QY 116 GlySerAlaAlaTyrMetAlaProGluValPheGluGlySerAsnTyrglyLysCys 135
DB 502 GGAACCTTCCCATGATGGCTCCAGAAATTATTCAGAGTCCCTGTCAGAAAACCTTGT 561
```

```
Qy 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 GACACATATCTCTATGTTGGTCTCTCGGAGATGCTAACAGGGAGGTCCTCTTTAAA 621

Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisnGlyThrArgPro 174
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 622 GGTITGGNAGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAAGAGATTA 675

Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 ACCATTCCAGCAGTTGCCCCAGAAGTTTTCGTAACCTGTTACATCAGTGTGGGAAGCT 735

Qy 195 AspProSerGlnArgProSerMetGluIleValLysIleMet 209
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 736 GATGCCAAGAAACGCCCATCATTTCAAGCAAAATCATTTCAATCCTG 780

RESULT 14
US-09-221-237-6
; Sequence 6, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-221-237-6

Alignment Scores:
Pred. No.: 6 56e-35 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 4 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-237-6 (1-1365)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 157 GAGCAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 216

Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 217 CTGAACCTCCCAACTATGTCATTGTTCACAGAATATGCTTCTCTGGGATCACTCTATGAT 276

Qy 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 277 TACATTAACTAACAGAGTGGAGATG-----GATATGATCACCATTATGACC 327

Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 328 TGGGCACTGATGTAGCCAAAGGAATGCATTATTACATATGAGGCTCTGTCAAGGTG 387

Qy 78 IleHisArgAspLeuLysProProAsnLeuValAlaGlyGlyThrValLeuLys 97
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 388 ATTACAGAGACCTCAAGTCAAGAAACGTTGTATAGTCTGATGGA---GTACTGAAG 444

Qy 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 445 ATCTGTGACTTTGGT---GCCTCTCGGTTCCATAACCAACACACACATGCTTGGTT 501
```

```
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 502 GGAACCTTTCCCATGGATGGCTCCAGAAGTTATCCAGAGTCTCCCTGTGTCAAAAACCTTGT 561

Qy 136 AspValPheSerTrpGlyIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 562 GACACATATCTCTATGTTGGTCTCTCTGGAGATGCTAACAGGGAGGTCCTCTTTAAA 621

Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 622 GGTITGGNAGA-----TTACAAGTAGCTTGGCTTGTAGTGGAAAAAACAAGAGATTA 675

Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 676 ACCATTCCAGCAGTTGCCCCAGAAGTTTTCGTAACCTGTTACATCAGTGTGGGAAGCT 735

Qy 195 AspProSerGlnArgProSerMetGluIleValLysIleMet 209
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 736 GATGCCAAGAAACGCCCATCATTTCAAGCAAAATCATTTCAATCCTG 780

RESULT 15
US-09-221-235-4
; Sequence 4, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-221-235-4

Alignment Scores:
Pred. No.: 1 27e-34 Length: 2120
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 3 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-221-235-4 (1-2120)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 203 GAGCAGAAATACTCAGTGTCTCAGTCACAGAAACATCATCCAGTTTATGGAGTAATT 262

Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 263 CTTGAACCTCCCAACTATGTCATTGTTCACAGAATATGCTTCTCTGGGATCACTCTATGAT 322

Qy 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 323 TACATTAACTAACAGAGTGGAGATG-----GATATGATCACCATTATGACC 373

Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 374 TGGGCACTGATGTAGCCAAAGGAATGCATTATTACATATGAGGCTCTGTCAAGGTG 433

Qy 78 IleHisArgAspLeuLysProProAsnLeuValAlaGlyGlyThrValLeuLys 97
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 434 ATTCAGAGACCTCAAGTCAAGAAACGTTGTATAGTCTGATGGA---GTACTGAAG 490
```

```

Qy 98 IleCysaspPheGlyThrAlaCysaspIleGlnThrHisMetThrAsnAsnLys----- 115
Db 491 ATCTGTGACTTTGGT---GCCCTCGGTTCCATAACCATACACACATGTCCTTGGTT 547
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
Db 548 GGAACCTTCCAGTGGCTCCAGAAAGTTATCCAGAGTCTCCCTGTGTCAAGAACTTGT 607
Qy 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
Db 608 GACACATATTCTTANGTGTGTGTTCTCTGGAGATGCTACAGAGGAGGTCCCTTTAA 667
Qy 156 GluIleGlyIleProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 668 GGTTCGAGAAGA-----TTACAGTAGCTTGCTGTAGTGAAAAAACAAGAGATT 721
Qy 175 ProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrpSerLys 194
Db 722 ACCATTCCAGGAGTTGGCCCGAAGTTTGCTGAACGTGTACATCAGTGTGGAGACT 781
Qy 195 AspProSerGlnArgProSerMetGluGluIleValLysIleMet 209
Db 782 GATGCCAGAAAGCGCATCATTTCAAGCAAAATCATTTCAATCCTG 826

```

Search completed: December 10, 2002, 05:36:02
 Job time : 74 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 10, 2002, 03:59:56 ; Search time 913 Seconds
(without alignments)
97.376 Million cell updates/sec

Title: US-09-830-144-2_COPY_76_303

Perfect score: 1252

Sequence: 1 VELRQLSRVNHNIKLYCA.....MTHLMRYPPGADPELOYPQC 228

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 350425 seqs, 194966369 residues

Total number of hits satisfying chosen parameters: 700850

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09830144/runat_04122002_141755_4398/app_query.fasta_1.391
-DB=Published Applications NA -QWTF=fastap -SUFFIX=p2n.rnmp -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09830144 @CGN 1.1 36 @runat 04122002_141755_4398
-NCPU=6 -ICPU=3 -NO XLPHY -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCT05_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1252	100.0	1788	9	US-10-158-895-14 Sequence 14, Appl
2	1252	100.0	2856	9	US-10-158-895-3 Sequence 3, Appl
3	392.5	31.3	3454	10	US-09-969-347-226 Sequence 226, App
4	371	29.6	1365	10	US-09-757-982-6 Sequence 6, Appl

5	371	29.6	2120	10	US-09-757-982-4	Sequence 4, Appl
6	355	28.4	3111	12	US-10-014-882-1	Sequence 1, Appl
7	355	28.4	3518	12	US-10-014-882-3	Sequence 3, Appl
8	347	27.7	1428	9	US-09-938-842A-882	Sequence 882, App
9	345.5	27.6	1662	9	US-09-938-842A-1014	Sequence 1014, Ap
10	338.5	27.0	2505	10	US-09-947-199-3	Sequence 3, Appl
11	338.5	27.0	3025	10	US-09-947-199-9	Sequence 9, Appl
12	335.5	26.8	2505	10	US-09-947-199-9	Sequence 9, Appl
13	335.5	26.8	3026	10	US-09-947-199-7	Sequence 7, Appl
14	328.5	26.2	1638	9	US-09-938-842A-903	Sequence 903, App
15	328.5	26.2	2211	9	US-09-938-842A-1577	Sequence 1577, Ap
16	313.5	25.0	3286	10	US-09-904-389-1	Sequence 1, Appl
17	304.5	24.3	2892	9	US-09-938-842A-1073	Sequence 1073, Ap
18	301.5	24.1	3663	9	US-09-919-172-84	Sequence 84, Appl
19	298.5	23.8	6383	9	US-09-954-531-405	Sequence 405, Appl
20	293.5	23.4	2574	10	US-09-735-103-2	Sequence 2, Appl
21	293.5	23.4	2574	12	US-10-045-428A-2	Sequence 2, Appl
22	289	23.1	1627	10	US-09-828-313-16	Sequence 16, Appl
23	288	23.0	3012	9	US-09-938-842A-1479	Sequence 1479, Ap
24	285	22.8	2661	9	US-09-938-842A-1085	Sequence 1085, Ap
25	278	22.2	3454	12	US-10-044-090-48	Sequence 48, Appl
26	278	22.2	3726	10	US-09-925-302-271	Sequence 271, App
27	271.5	21.7	2719	10	US-09-917-800A-1405	Sequence 1405, Ap
28	270.5	21.6	1911	10	US-09-917-800A-1611	Sequence 1611, Ap
29	270	21.6	2184	10	US-09-728-952-82	Sequence 82, Appl
30	269.5	21.5	2000	10	US-09-977-269-1	Sequence 1, Appl
31	268.5	21.4	2869	10	US-09-731-231A-1	Sequence 1, Appl
32	268	21.4	1052	9	US-09-938-842A-553	Sequence 553, App
33	268	21.4	3791	10	US-09-757-100B-1	Sequence 1, Appl
34	266.5	21.3	1296	10	US-09-906-397-2	Sequence 2, Appl
35	266.5	21.3	2016	10	US-09-758-003-1	Sequence 1, Appl
36	266.5	21.3	2322	10	US-09-906-397-1	Sequence 1, Appl
37	266.5	21.3	2322	10	US-09-906-397-5	Sequence 5, Appl
38	266	21.2	4268	10	US-09-954-456-293	Sequence 293, App
39	266	21.2	4268	10	US-09-954-456-1599	Sequence 1599, Ap
40	265.5	21.2	3911	12	US-10-000-864-1	Sequence 1, Appl
41	264.5	21.1	2015	10	US-09-954-456-1983	Sequence 1983, Ap
42	264	21.1	2782	9	US-10-108-605-250	Sequence 250, App
43	262.5	21.0	3260	10	US-09-858-754-1	Sequence 1, Appl
44	262.5	21.0	5253	10	US-09-858-754-13	Sequence 13, Appl
45	262.5	21.0	5253	12	US-10-000-864-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-10-158-895-14
; Sequence 14, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/Jp98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 1788
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(1776)

US-10-158-895-14

Alignment Scores:

Pred. No.:	1,33e-138	Length:	1788
Score:	1252.00	Matches:	228
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2_COPY_76_303 (1-228) x US-10-158-895-14 (1-1788)

```

QY 1 VALGILUEUARGILNUSERARGVALASNHISPROASNIIEVALYSEUURYGIYALA 20
    |||
DB 232 GTAGAGCTTCGGCAGTTATCCCGTGAACCACTCTAATATTGTTAAAGCTTATGAGACC 291

QY 21 CYSLEUASNPROVALCYSLEUVALMETGIUTYRALAGIUGIYGLYSEIEUTYRASNVAL 40
    |||
DB 292 TGGTTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATG 351

QY 41 LEUHSIGIYALAGIUPROLEUPROTYRYTYRTHRALAALAHISALAMESERTPYCYLEU 60
    |||
DB 352 CTGCATGTGCTGACCACTATGTCATATTACTGTGCTGCCACGCAATGATGTTGTTTA 411

QY 61 GINCYSEERGINGLYVALALATYRLEUHSISSEMERGLNPROLYSALALEUILEHISARG 80
    |||
DB 412 CAGTGTTCACAGAGAGCTGCTTATCTTCAAGCATGCAACCCAAAGCGCTTATTCACAGG 471

QY 81 ASPLEUENLSPROASNEULEUENLEUVALAAGIYGLYTHRVALLLEUYSIIECYASAP 100
    |||
DB 472 GACCTGAACCAACCAACTTACTGCTGTGTGACGGGGGACAGTTCTAATAATTTGTGAT 531

QY 101 PHEGITYTHRALACYASAPIIIEGINTHRHISMERTHASNSNLYSGIYSEIALAATRP 120
    |||
DB 532 TTTGGTACAGCCCTGTGACATTCAGACACATGACCAATTAACAAGGGGAGTCTGCTGG 591

QY 121 METALPROGLIUVAPHEGIUGIYSEASNTYRSEGLIYSCYASAPVALPHESERTRP 140
    |||
DB 592 ATGGCACCTGGAAGTTTTTGAAGGTAGTAATTACAGTGAAGAAAATGTGACCTTCCAGCTGG 651

QY 141 GLYILEILEUTRPGIUVALLIETHRARGLYSPROPHASPGIUIIEGLIYGLYPRO 160
    |||
DB 652 GGTATTATTTCTTTGGGAGATGATACCGCTCGGAACCCCTTGATGATGATGTTGGGCCCA 711

QY 161 ALAPHEARGIIEWECTYRALAVALHISASNGIYTHRARGPROLEUIIELYSASNLEU 180
    |||
DB 712 GCTTCCGAATCATGTGGGCTGTTCATTAATGTACTGACCAACCACTGATTAATAATTTTA 771

QY 181 PROLYSPROIIIEGLIUSERLEUMETTHRARGCYSTRPSELYSAPPROSERGINARPRO 200
    |||
DB 772 CCTAAGCCCATTTGAGAGCTGATGACTCGTTGTTGGTCTAAGATCTTCCACGCGCCCT 831

QY 201 SEMETGIUGIUIIEVALYSLIEMETTHRHISLEUMETARGTYRPHETPROGLIYALAASP 220
    |||
DB 832 TCATGAGAGAAATTGTGAAAATAATGACCACTTATGCGGTACTTCCAGAGAGAGAT 891

QY 221 GIUPROLEUGINTYRPROCYSGIN 228
    |||
DB 892 GAGCCATTACAGTATCTTGTCTAG 915

RESULT 2
US-10-158-895-3
; Sequence 3, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF- $\beta$  INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11

```

```

; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2656
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (183)..(1919)
US-10-158-895-3

```

Alignment Scores:

Pred. No.:	2.37e-138	Length:	2656
Score:	1252.00	Matches:	228
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-830-144-2_COPY_76_303 (1-228) x US-10-158-895-3 (1-2656)

```

QY 1 VALGILUEUARGILNUSERARGVALASNHISPROASNIIEVALYSEUURYGIYALA 20
    |||
DB 408 GTAAGAGCTTCGGCAGTTATCCCGTGAACCACTCTAATATTGTTAAAGCTTATGAGACC 467

QY 21 CYSLEUASNPROVALCYSLEUVALMETGIUTYRALAGIUGIYGLYSEIEUTYRASNVAL 40
    |||
DB 468 TGGTTGAATCCAGTGTGCTTGTGATGGAATATGCTGAAGGGGGCTCTTATATATG 527

QY 41 LEUHSIGIYALAGIUPROLEUPROTYRYTYRTHRALAALAHISALAMESERTPYCYLEU 60
    |||
DB 528 CTGCATGTGCTGACCACTATGTCATATTACTGTGCCACCAACCAATGATGTTGTTTA 587

QY 61 GINCYSEERGINGLYVALALATYRLEUHSISSEMERGLNPROLYSALALEUILEHISARG 80
    |||
DB 588 CAGTGTTCACAGAGAGCTGCTTATCTTCAAGCATGCAACCCAAAGCGCTTATTCACAGG 647

QY 81 ASPLEUENLSPROASNEULEUENLEUVALAAGIYGLYTHRVALLLEUYSIIECYASAP 100
    |||
DB 648 GACCTGAACCAACCAACTTACTGCTGTGTGACGGGGGACAGTTCTAATAATTTGTGAT 707

QY 101 PHEGITYTHRALACYASAPIIIEGINTHRHISMERTHASNSNLYSGIYSEIALAATRP 120
    |||
DB 708 TTTGGTACAGCTGTGACATTCAGACACATGACCAATTAACAAGGGGAGTCTGCTGG 767

QY 121 METALPROGLIUVAPHEGIUGIYSEASNTYRSEGLIYSCYASAPVALPHESERTRP 140
    |||
DB 768 ATGGCACCTGGAAGTTTTTGAAGGTAGTAATTACAGTGAAGAAAATGTGACGTTCACTGG 827

QY 141 GLYILEILEUTRPGIUVALLIETHRARGLYSPROPHASPGIUIIEGLIYGLYPRO 160
    |||
DB 828 GGTATTATTTCTTTGGGAGATGATACCGCTCGGAACCCCTTGATGATGATGTTGGGCCCA 887

QY 161 ALAPHEARGIIEWECTYRALAVALHISASNGIYTHRARGPROLEUIIELYSASNLEU 180
    |||
DB 888 GCTTCCGAATCATGTGGGCTGTTCATTAATGTACTGACCAACCACTGATTAATAATTTTA 947

QY 181 PROLYSPROIIIEGLIUSERLEUMETTHRARGCYSTRPSELYSAPPROSERGINARPRO 200
    |||
DB 948 CCTAAGCCCATTTGAGAGCTGATGACTCGTTGTTGGTCTAAGATCTTCCACGCGCCCT 1007

QY 201 SEMETGIUGIUIIEVALYSLIEMETTHRHISLEUMETARGTYRPHETPROGLIYALAASP 220
    |||
DB 1008 TCATGAGAGAAATTGTGAAAATAATGACCACTTATGCGGTACTTCCAGAGAGAGAT 1067

QY 221 GIUPROLEUGINTYRPROCYSGIN 228
    |||
DB 1068 GAGCCATTACAGTATCTTGTCTAG 1091

```

RESULT 3

US-09-969-347-226
; Sequence 226, Application US/09969347
; Patent No. US20020115085A1
; GENERAL INFORMATION:
; APPLICANT: Ebner, Reinhard
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Signatu
; FILE REFERENCE: 689290-69
; CURRENT APPLICATION NUMBER: US/09/969,347
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US/60/237,598
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: US/60/237,604
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 318
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 226
; LENGTH: 3454
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-969-347-226

Alignment Scores:
Pred. No.: 1.7e-36 Length: 3454
Score: 392.50 Matches: 88
Percent Similarity: 56.68% Conservative: 35
Best Local Similarity: 40.55% Mismatches: 77
Query Match: 31.35% Indels: 17
DB: 10 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-09-969-347-226 (1-3454)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db 721 GAGCCCGGCTCTTTGAGCGCCTTGACGACCCCAACATTAATCCCTTAGGGCGGCTGC 780
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlySerLeuTyrAsn 39
Db 781 CTCAACCCCCACACCTCTCGCTAGTAGTGAGTATGCCCGGGTGTGCACCTGACGAGG 840
Qy 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrpCys 59
Db 841 GTGTGTCAGTGC CGCGGGTGCACCTCACGTG-----CTGGTCAACTGGGCT 888
Qy 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuIleHis 79
Db 889 GTGACGTGCGCGGGGCATGAACCTACCAATGATGCCCTCTGCCCATCATCCAC 948
Qy 80 ArgAspLeuLysProProAsnLeuLeuVal-----AlaGly 92
Db 949 CGGGACCTCAAGTCCATCAACATCCTGATCCTGGAGGCCATCGAAGCACCACTCGCA 1008
Qy 93 GlyThrValLeuLysIleCysAspPheGlyThrAlaCysAsp---lleGlnThrHisMet 111
Db 1009 GACACGGTGTCAAGATCACGGACTTCGGGCTCGCCCGGAGTGGGCACCAAGACCAAG 1068
Qy 112 ThrAsnAsnLysGlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyr 131
Db 1069 ATGAGCGCTCGGGGACCTACGCGCTGGATGCGCGCGAGGTTATCCGTCTCTCCCTCTTC 1128
Qy 132 SerGluLysCysAspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArg 151
Db 1129 TCCAAAGACGAGTATGCTCGAGCTTCGGGGTGTGCTGTGGAGGTGTGACGGGGGAG 1188
Qy 152 LysProPheAspGluIleGlyProAlaPheArgIleMetTrpAlaVal---HisAsn 170
Db 1189 GTCCCTTACCGTGAGATC-----GACGCTTGGCGGTGCGGTATGGCGGTGATGAAT 1242
Qy 171 GlyThrArgProProLeuIleLysAsnLeuProLysProIleGluSerLeuMetThrArg 190
Db 1243 AAGCTGACGCTGCCCATTCCTCCACGTGCGCCCGAGCCCTTTGCCCGCCTCTCTGGAGAA 1302

Qy 191 CysTrpSerLysAspProSerGlnArgProSerMetGluGluIleValLys 207
Db 1303 TGCTGGACCCAGACCCCGCCGCGCCAGATTTCGGTAGCATCTTTGAAG 1353

RESULT 4

US-09-757-982-6
; Sequence 6, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; PRIOR FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1365
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1365)
US-09-757-982-6
Alignment Scores:
Pred. No.: 1.54e-34 Length: 1365
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-757-982-6 (1-1365)

Qy 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValIysLeuTyrGlyAlaCys 21
Db 157 GAGCGAGAAATACTCAGTGTCTCAGTCAGACAGAAACATCATCCAGTTTATGGAGTAAT 216
Qy 22 LeuAsnPro-----ValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyrAsn 39
Db 217 CTTGAACCTCCCAACTATGGCAATGTGCAGAAATATGTTCTCTGGGATCACTCTATGAT 276
Qy 40 ValLeuHisGly-----AlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSer 57
Db 277 TACATTAACAGTAACAGAAAGTGAGGAGATG-----GATATGGATCACTATTATGACC 327
Qy 58 TrpCysLeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeu 77
Db 328 TGGGCCACTGATGTAGCCAAAGGAATGCATTATTATACATATGGAGGCTCCTGTCAAGGTG 387
Qy 78 IleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 388 ATTCACAGAGACCTCAAGTCAAGAAACGTTGTTATAGTGTCTGATGGA--GTACTGAAG 444
Qy 98 IleCysAspPheGlyThrAlaCysAspIleGlnThrHisMetThrAsnAsnLys----- 115
Db 445 ATCTGTGACTTTGGT--GCCTCTCGGTTCCATAACCATACACACATGTCCTTGGTT 501
Qy 116 GlySerAlaAlaTrpMetAlaProGluValPheGluGlySerAsnTyrSerGluLysCys 135
Db 502 GGAACCTTCCCATGGATGGCTCCAGAAAGTATCCAGAGTCTCCCTGTGTGCAAACTTGT 561
Qy 136 AspValPheSerTrpGlyIleIleLeuTrpGluValIleThrArgArgLysProPheAsp 155
Db 562 GACACATATTCCTATGGTGTGGTTCTCTGGGAGATGCTAAACAAGGAGGTCCTCCCTTTAAA 621
Qy 156 GluIleGlyGlyProAlaPheArgIleMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 622 GGTTTGGAAGGA-----TTACAAGTAGCTTGGCTTTAGTGGAAAAAACAACAGAGATTA 675

QY 175 ProLeu11eLysasnLeuProLysPro11eGluSerLeuMetThrArgCysTrpSerLys 194
Db 676 ACCATTCCAAAGCAGTGGCCCCAGAAAGTTTGTCTGAACCTGTACATCACTGTGGGAAGCT 735
QY 195 AspProSerGlnArgProSerMetGluGlu11eValLys11eMet 209
Db 736 GATGCCAAGAAAGGCCATCATTCACAGCAATCATTCATCTG 780

RESULT 5

US-09-757-982-4
; Sequence 4, Application US/09757982
; Patent No. US20020094559A1
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAAP-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/757,982
; CURRENT FILING DATE: 2001-01-10
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 2120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (47)..(1411)
US-09-757-982-4

Alignment Scores:
Pred. No.: 2,92e-34 Length: 2120
Score: 371.00 Matches: 81
Percent Similarity: 56.28% Conservative: 40
Best Local Similarity: 37.67% Mismatches: 80
Query Match: 29.63% Indels: 14
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-757-982-4 (1-2120)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsn11eValLysLeuTrg1yAlaCys 21
Db 203 GAGGCGAAATATCTCAGTCTCCTCAGTCAACAGAAATCATCAGTTTATGAGTAATT 262
QY 22 LeuAsnPro-----ValCysLeuValMetGluTrgAlaGluGlyGlySerLeuTrgAsn 39
Db 263 CTTGAACTCCCAACTAGTATGAGCATGTGCACAGAAATATGCTTCTGGGATCATCTATGAT 322
QY 40 ValLeuHisGly-----AlaGluProLeuProTrgTrpThrAlaAlaHisAlaMetSer 57
Db 323 TACATTAAACAGTAACAGAGTGGAGATG-----GATATGAGTACATTTATGACC 373
QY 58 TrpCysLeuGlnCysSerGlnGlyValAlaTrpLeuHisSerMetGlnProLysAlaLeu 77
Db 374 TGGGCCCATGATGTACCAAGGAATGCAATTATTAATATGAGAGCTCTCTCAAGGTG 433
QY 78 11eHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 434 ATTCACAGAGAGCTCAAGTCCAGAAACGTTTATATAGCGCTGAGTGA---CTACTGAAG 490
QY 98 11eCysAspPheGlyThrAlaCysAsp11eGlnThrHisMetThrAsnAsnLys----- 115
Db 491 ATCTGTGACTTGGT---CCCTCTCGGTTCCATTAACCATTAACAACACATGTCTTGGTT 547
QY 116 GlySerAlaAlaTrpMetAlaProGluValAlaPheGluGlySerAsnTrpSerGlnLysCys 135
Db 548 GGAACCTTCCCAAGTATGCTCCAGAAATTCACAGAGTCTCCCTGTGCAGAAACCTTGT 607
QY 136 AspValPheSerTrpGlyLys11eLeuTrpGluValAla11eThrArgArgLysProPheAsp 155
Db 608 GACACATATTCCTATGTGTGTGTTCTCTGGAGATGCTTAACAAGGAGGTCCCTTTAA 667

QY 156 Glu11eGlyGlyProAlaPheArg11eMetTrp---AlaValHisAsnGlyThrArgPro 174
Db 668 GGTTCGAGAGA-----TTACAAAGTATGCTTGCTGTGATGAAAAAAGAGATTA 721
QY 175 ProLeu11eLysasnLeuProLysPro11eGluSerLeuMetThrArgCysTrpSerLys 194
Db 722 ACCATTCCAAAGCAGTGGCCCCAGAAAGTTTGTCTGAACCTGTACATCACTGTGGGAAGCT 781

QY 195 AspProSerGlnArgProSerMetGluGlu11eValLys11eMet 209
Db 782 GATGCCAAGAAAGGCCATCATTCACAGCAATCATTCATCTG 826

RESULT 6

US-10-014-882-1
; Sequence 1, Application US/10014882
; Patent No. US20020107384A1
; GENERAL INFORMATION:
; APPLICANT: Hu, Yi
; APPLICANT: Kieke, James
; APPLICANT: Donoho, Gregory
; TITLE OF INVENTION: No. US20020107384A1 Human Kinase and Polynucleotides Encoding t
; FILE REFERENCE: LEX-0279-USA
; CURRENT APPLICATION NUMBER: US/10/014,882
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,744
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3111
; TYPE: DNA
; ORGANISM: homo sapiens
US-10-014-882-1

Alignment Scores:
Pred. No.: 3,99e-32 Length: 3111
Score: 355.00 Matches: 83
Percent Similarity: 53.48% Conservative: 40
Best Local Similarity: 36.09% Mismatches: 85
Query Match: 28.35% Indels: 22
DB: 12 Gaps: 6

US-09-830-144-2_COPY_76_303 (1-228) x US-10-014-882-1 (1-3111)

QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsn11eValLysLeuTrg1yAlaCys 21
Db 511 GAGGCTCGGCTCTTCGATGCTGCGGACCCCAACATCATCGAGCTGGCGCGCTGCTG 570
QY 22 LeuAsnPro-----ValCysLeuValMetGluTrgAlaGluGlyGlySerLeuTrgAsn 39
Db 571 CTGCAGAGCGGACCTCTCTGCTGCTGAGATTGCGCGCGGAGCGCTCAACCGA 630
QY 40 ValLeuHisGlyAlaGluProLeuProTrgTrpThrAla----- 52
Db 631 GCGGTGCGCTGCAAGCGCCGCCGCGGACCCGCGCGCGCGCGCGCGCGCGCGCGCG 690
QY 53 -----AlaHisAlaMet---SerTrpCysLeuGlnCysSerGlnGlyValAlaTrp 68
Db 691 CGCATCCCTCGGACCGTGTGTCACATGCGCGCGCTGCACATGATACGAGGCGCATGCTTC 750
QY 69 LeuHisSerMetGlnProLysAlaLeu11eHisArgAspLeuLysProProAsnLeuLeu 88
Db 751 CTGCATAGAGAGGCTTGTGCTCCCATCTGACACCGGAGCTCAAGTCCAGCAACATTTTG 810
QY 89 LeuValAlaGly-----GlyThrValLeuLys11eCysAspPhe 101
Db 811 CTACTTGAAGAATAGAACTGATGACATGTGCATTAATAAATTGAAAGTTACAGATTTT 870
QY 102 GlyThrAlaCysAsp---11eGlnThrHisMetThrAsnAsnLysGlySerAlaAlaTrp 120
Db 871 GGGTTGGCGAGGAGAGGACAGAGACCAACAAATAGAGACAGAGACGATATGCTGG 930
QY 121 MetAlaProGluValPheGluGlySerAsnTrpSerGlnLysCysAspValPheSerTrp 140

[illegible]

Db 958 GAAGTATATATATAGAGGAAAGTTGGCATATAAATGTTGCAGTTTCATTTGGTCATGT 1017

QY 22 LeuAsnPro-----ValCysLeuValMetGluYrAlaGluGlySerLeuYrAsn 39

Db 1018 ACAAGCATCACCAAACTTCGATTTGTGCAAGAGTTCAAGTTCGCGGGAGAGATTATATGAT 1077

QY 40 ValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaHisAlaMetSerTyrCys 59

Db 1078 TTCCTTCAACAAACAAAGGGTT-----TTTAAATTCATCTTTGCTCAAGTGGCA 1131

QY 60 LeuGlnCysSerGlnGlyValAlaTyrLeuHisSerMetGlnProLysAlaLeuHis 79

Db 1132 CTCGACGTTCCAAAGGAATGATATATCTGCATCA-----AACCAATATTATTCAT 1182

QY 80 ATGAspLeuLysProProAsnLeuLeuValAlaGlyThrValLeuLysIleCys 99

Db 1183 AAGAGACTTAAAGACTGCTAACTCTTCTTATG--GACGAAACATGAAGTTGTCAAAAGTTGCC 1239

QY 100 AspPheGlyThrAlaCysAspIleGlnThrHis-----MetThrAsnLysGly 116

Db 1240 GATTTTGGTGTTGCC---AGAGTCGACACTGACGACAGGGTTATGACAGGGAAACAGGG 1296

QY 117 SerAlaAlaTyrPheAlaProGluValPheGluGlySerAsnTyrSerGluLysCysAsp 136

Db 1297 ACATACCATGATGATGGCTTCACAGACTCATTTAGACAAACCTTATGATCATCAGGGCAGAT 1356

QY 137 ValPheSerTyrGlyIleIleLeuTyrGluValIleThrArgArgLysProPheAspGlu 156

Db 1357 GCTTCACAGTACGACGATTTGTCTGTGGAACTTTGACTGGGAACCTCCCATATTTCTTAC 1416

QY 157 IleGlyGlyProAlaPheArgIleMetTyrAlaValHisAsnGlyThrArgProProLeu 176

Db 1417 TTG---ACTCACACTGCACAGCTGCTGTGGCGTTGTCCAAAAGGACCTTAGACCAAAATT 1473

QY 177 IleLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTyrSerLysAspPro 196

Db 1474 CCMAAGGAAACACACCCAAACTACTGAACCTTCTTGAAGAAATGCTGGCAGCAAGACCCA 1533

QY 197 SerGlnArgProSerMetGluGluIleValLysIleMetThrHisLeuMetArg----- 214

Db 1534 GCTCAATAGACCAACATTTTGCAGAAATCATGAAATGCTTAAACCAACTAATCCGCGAGTAA 1593

QY 215 -----TyrPheProGly 218

Db 1594 ATTGATTTATCATTCATTAAGATAAACATGCTGGTTACTTTTCAGGC 1641

RESULT 10

US-09-947-199-3

Sequence 3, Application US/09947199

Patent No. US20020127684A1

GENERAL INFORMATION:

APPLICANT: Raju, Jeyaseelan

TITLE OF INVENTION: NOVEL CARC PROTEIN AND NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: NNI-068CP2

CURRENT APPLICATION NUMBER: US/09/947,199

CURRENT FILING DATE: 2001-09-05

PRIOR APPLICATION NUMBER: 60/111,938

PRIOR FILING DATE: 1998-12-11

PRIOR APPLICATION NUMBER: 09/291,839

PRIOR FILING DATE: 1999-04-14

PRIOR APPLICATION NUMBER: 09/458,457

PRIOR FILING DATE: 1999-12-10

NUMBER OF SEQ ID NOS: 9

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 3

LENGTH: 2505

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (1)..(2505)

US-09-947-199-3

```
Alignment Scores:
Pred. No.: 2,61e-30 Length: 2505
Score: 338.50 Matches: 84
Percent Similarity: 59.07% Conservative: 43
Best Local Similarity: 39.07% Mismatches: 72
Query Match: 27.04% Indels: 17
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-3 (1-2505)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 1525 GAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGGTGCTTGC 1584
QY 22 LeuAsn-----ProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 38
Db 1585 TTGAATGATCCAGCCAGTTGCCATTGTCTCACTCAATACATATCAGGGGGTCTCTGTTT 1644
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 58
Db 1645 TCCCTCTTCATGA-GCAGAAGAGGATTCTTGATTT-----GCAGTCTAAATAATTAT 1697
QY 59 CysLeuGlnCysSerGln-GlyValAlaTyrLeuHisSerMet---GlnProLysAlaLe 77
Db 1698 TGCAGTAGATGTTGCCAAGGCGATGGAGTACCTTCAACACTGACACAGCCA-----AT 1751
QY 77 uIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 1752 TATACATCGTGACTTGAACAGTCAATATCTCTCTATGAGGATGGGATCGCTGTG-- 1809
QY 97 sIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMetThrAs 113
Db 1810 -GTGGCAGATTTTGGAGAAATCAAGATTCTACAGTCTCTGGATGAAGACAAATGACAAA 1868
QY 113 nAsnLysGlySerAlaAlaTrpMetAlaProGluValPhe---GluGlySerAsnTyrSe 132
Db 1869 ACAACCTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCCAGCAGTGCACTCGGTAC 1928
QY 132 rGluLysCysAspValPheSerTrpGlyIleLeuTyrGluValIleThrArgArgly 152
Db 1929 CATCAAGCAGATGCTTTCAGTATGCTCTGTGCTGTGGAAATCTCTACCTGGCGAAAT 1988
QY 152 sProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyTh 172
Db 1989 TCCATTGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGCTTACCACCAC---AT 2045
QY 172 rArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTr 192
Db 2046 CAGACCTCCCATGGCTATTTCATTCCTCAAGCCCATATCATCTCTGCTGATACGAGGGTG 2105
QY 192 pSerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2106 GAACGCATGCTCTGAAGAGACCCGAATTTTCTGAAGTTGTC 2148

RESULT 11
US-09-947-199-1
; Sequence 1, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 3025
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (48)..(2552)
US-09-947-199-1
Alignment Scores:
Pred. No.: 3,43e-30 Length: 3025
Score: 338.50 Matches: 84
Percent Similarity: 59.07% Conservative: 43
Best Local Similarity: 39.07% Mismatches: 72
Query Match: 27.04% Indels: 17
DB: 10 Gaps: 8

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-1 (1-3025)
QY 2 GluLeuArgGlnLeuSerArgValAsnHisProAsnIleValLysLeuTyrGlyAlaCys 21
Db 1572 GAGGTGTCATCTCTGCCAGCTCAATCATCCCTGCGTAATTCAGTTGTGGGTGCTTGC 1631
QY 22 LeuAsn-----ProValCysLeuValMetGluTyrAlaGluGlyGlySerLeuTyr 38
Db 1632 TTGAATGATCCAGCCAGTTTGCATTCCTCACTCAATACATATCAGGGGGTCTCTGTTT 1691
QY 39 AsnValLeuHisGlyAlaGluProLeuProTyrTyrThrAlaAlaHisAlaMetSerTrp 58
Db 1692 TCCCTCTTCATGA-GCAGAAGAGGATTCTTGATTT-----GCAGTCTAAATAATTAT 1744
QY 59 CysLeuGlnCysSerGln-GlyValAlaTyrLeuHisSerMet---GlnProLysAlaLe 77
Db 1745 TGCAGTAGATGTTGCCAAGGCGATGGAGTACCTTCAACACTGACACAGCCA-----AT 1798
QY 77 uIleHisArgAspLeuLysProProAsnLeuLeuValAlaGlyGlyThrValLeuLys 97
Db 1799 TATACATCGTGACTTGAACAGTCAATATTTCTCTCTATGAGGATGGGATCGCTGTG-- 1856
QY 97 sIleCysAspPheGlyThrAlaCysAspIleGlnThr-----HisMetThrAs 113
Db 1857 -GTGGCAGATTTTGGAGAAATCAAGATTCTTACAGTCTCTGGATGAAGACAAATGACAAA 1915
QY 113 nAsnLysGlySerAlaAlaTrpMetAlaProGluValPhe---GluGlySerAsnTyrSe 132
Db 1916 ACAACCTGGGAACCTCCGTTGGATGGCTCTGAGGTGTTTCCAGCAGTGCACTCGGTAC 1975
QY 132 rGluLysCysAspValPheSerTrpGlyIleLeuTyrGluValIleThrArgArgly 152
Db 1976 CATCAAGCAGATGCTTTCAGTATGCTCTGTGCTGTGGGAAATCTCTCACTGGCGAAAT 2035
QY 152 sProPheAspGluIleGlyGlyProAlaPheArgIleMetTrpAlaValHisAsnGlyTh 172
Db 2036 TCCATTGCTCATCTCAAGCCAGCGGCTGGCGCAGCAGACATGCTTACCACCAC---AT 2092
QY 172 rArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTr 192
Db 2093 CAGACCTCCCATGGCTATTTCATTCCTCAAGCCCATATCATCTCTGCTGATACGAGGGTG 2152
QY 192 pSerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2153 GAACGCATGCTCTGAAGAGACCCGAATTTTCTGAAGTTGTC 2195

RESULT 12
US-09-947-199-9
; Sequence 9, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: THEREFOR
; FILE REFERENCE: MNI-068CP2
```

```

; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2505)
; US-09-947-199-9
```

```

Alignment Scores:
Pred. No.: 5,9e-30 Length: 2505
Score: 335.50 Matches: 80
Percent Similarity: 56.54% Conservative: 41
Best Local Similarity: 37.38% Mismatches: 78
Query Match: 26.80% Indels: 15
DB: 10 Gaps: 8
```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-9 (1-2505)

```

QY 2 GtluLeuArgGlnLeuSerArgValaAenHisProAsnIleValLysLeuTyrgIalAcys 21
Db 1525 GAGGTGTCAATTCCTGCGACGCTCAACACCCCTGCGTGTTCAGTTGTGGTCCCTGC 1584
QY 22 Leu-----AsnProValCysLeuValMetGluTyrgIalGluGlyGlySerLeuTy 38
Db 1585 CTGGATGACCCCGACGATTGCGATTCCTGCTCACTCACTCAATTCAGAGGCTCCCTGTT 1644
QY 39 AsnValLeuHisGlyValaGluProLeuProTyrgTyrgThraIalaHisIalaMetSerTrp 58
Db 1645 TCCCTGCTTCAT-----GAAACAAGAGAAATTCCTGACTTCAGTCTTAATTAATTCATT 1698
QY 59 CysLeuGlnCysSerGlnGlyValaIaTyrgLeuHisSerMet---GlnProLysAlaLeu 77
Db 1699 GCGGTAGACGTTGCCAAGGCGCATGAGTACCTGCACAGCTTGAACCCAGCA-----ATC 1752
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValaIaGlyGlyThraValLeuLys 97
Db 1753 ATACACCGGACCTGGAACAGCCACAAATATTCGCTCTATGAGATGCGCATGCTGTG--- 1809
QY 98 IleCysAspPheGlyThraIaCysAspIleGlnThr-----HisMetThrAsn 113
Db 1810 GTGGCAGATTTGGAGAAATCAAGATTCGTCAGTCCCTGATGAAGACAAATGACAAAG 1869
QY 114 AsnLysGlySerAlaIaIaTrpMetAlaProGluValaPhe---GluGlySerAsnTySer 132
Db 1870 CAGCCAGGAAACCTGCGCTGATGAGCGCCCTGAGGTGTTCACACAGGCGACGAGATACCC 1929
QY 133 GluLysCysAspValaPheSerTrpGlyIleIleLeuTyrgIuValaIleThrArgArgLys 152
Db 1930 ATCAAGCGCTGATCTTCACTTAATCTCCCTGTGTGTGGAGCTCTCTCACTGAGAAATTT 1989
QY 153 ProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaValHisAsnGlyThr 172
Db 1990 CCATTGCTCATCTCAAGCAGCGCTGCAGACAGATATGCGGATACACAC---ATC 2046
QY 173 ArgProProLeuLysAsnLeuProLysProIleGluSerLeuMetThrArgCysTrp 192
Db 2047 AGACCCCGCATGCGCTATTCATCCCAAGCCCATCTCATCCCTGCTGTAGTACGGGGCTGG 2106
QY 193 SerLysAspProSerGlnArgProSerMetGluGluIleVal 206
Db 2107 AATGATGTCCTGGAAGACGACCCAGAGTTCTCTGAAGTGCTT 2148
```

```

RESULT 13
US-09-947-199-7
; Sequence 7, Application US/09947199
; Patent No. US20020127684A1
; GENERAL INFORMATION:
; APPLICANT: Raju, Jeyaseelan
; TITLE OF INVENTION: NOVEL CARP PROTEIN AND NUCLEIC ACID MOLECULES AND USES
; FILE OF INVENTION: THEREFOR
; FILE REFERENCE: MNI-068CP2
; CURRENT APPLICATION NUMBER: US/09/947,199
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: 60/111,938
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 09/291,839
; PRIOR FILING DATE: 1999-04-14
; PRIOR APPLICATION NUMBER: 09/458,457
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 3026
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(2565)
; US-09-947-199-7
```

```

Alignment Scores:
Pred. No.: 7,77e-30 Length: 3026
Score: 335.50 Matches: 80
Percent Similarity: 56.54% Conservative: 41
Best Local Similarity: 37.38% Mismatches: 78
Query Match: 26.80% Indels: 15
DB: 10 Gaps: 8
```

US-09-830-144-2_COPY_76_303 (1-228) x US-09-947-199-7 (1-3026)

```

QY 2 GtluLeuArgGlnLeuSerArgValaAenHisProAsnIleValLysLeuTyrgIalAcys 21
Db 1585 GAGGTGTCAATTCCTGCGACGCTCAACACCCCTGCGTGTTCAGTTGTGGTCCCTGC 1644
QY 22 Leu-----AsnProValCysLeuValMetGluTyrgIalGluGlyGlySerLeuTy 38
Db 1645 CTGGATGACCCCGACGATTGCGATTCCTGCTCACTCACTCAATTCAGAGGCTCCCTGTT 1704
QY 39 AsnValLeuHisGlyValaGluProLeuProTyrgTyrgThraIalaHisIalaMetSerTrp 58
Db 1705 TCCCTGCTTCAT-----GAAACAAGAGAAATTCCTGACTTCAGTCTTAATTAATTCATT 1758
QY 59 CysLeuGlnCysSerGlnGlyValaIaTyrgLeuHisSerMet---GlnProLysAlaLeu 77
Db 1759 GCGGTAGACGTTGCCAAGGCGCATGAGTACCTGCACAGCTTGAACCCAGCA-----ATC 1812
QY 78 IleHisArgAspLeuLysProProAsnLeuLeuValaIaGlyGlyThraValLeuLys 97
Db 1813 ATACACCGGACCTGGAACAGCCACAAATATTCGCTCTATGAGATGCGCATGCTGTG--- 1869
QY 98 IleCysAspPheGlyThraIaCysAspIleGlnThr-----HisMetThrAsn 113
Db 1870 GTGGCAGATTTGGAGAAATCAAGATTCGTCAGTCCCTGATGAAGACAAATGACAAAG 1929
QY 114 AsnLysGlySerAlaIaIaTrpMetAlaProGluValaPhe---GluGlySerAsnTySer 132
Db 1930 CAGCCAGGAAACCTGCGCTGATGAGCGCCCTGAGGTGTTCACACAGGCGACGAGATACACC 1989
QY 133 GluLysCysAspValaPheSerTrpGlyIleIleLeuTyrgIuValaIleThrArgArgLys 152
Db 1990 ATCAAGCGCTGATCTTCACTTAATCTCCCTGTGTGTGGAGCTCTCTCACTGAGAAATTT 2049
QY 153 ProPheAspGluIleGlyLysProAlaPheArgIleMetTrpAlaValHisAsnGlyThr 172
Db 2050 CCATTGCTCATCTCAAGCAGCGCTGCAGACAGATATGCGGATACAC---ATC 2106
```


GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 18:31:39 ; Search time 53.0526 Seconds
(without alignments)
1265.881 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLSQSQPSWTDD.....AEFYRLNSVDHGEQSVVTAP 504

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- A_Geneseq_101002:*
- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
 - 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
 - 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
 - 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
 - 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
 - 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
 - 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
 - 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
 - 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
 - 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
 - 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
 - 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
 - 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
 - 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
 - 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
 - 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
 - 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
 - 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
 - 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
 - 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
 - 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
 - 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
 - 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2580	100.0	504	18 AAW26706	Human TAB1 (TAK1 b
2	2580	100.0	504	20 AAY09541	Human TAB1 protein
3	2580	100.0	504	21 AAY91001	Human TAB-1 protei
4	2580	100.0	504	21 AAY59450	Human TAB1 protein
5	2580	100.0	513	20 AAY09550	Human TAB1 protein
6	2580	100.0	517	20 AAY09546	Human TAB1-FLAG pr
7	2575	99.8	504	18 AAW26707	Human TAB1 (TAK1 b
8	409	15.9	84	21 AAB56692	Human prostate can
9	365	14.1	70	22 ABB28132	Human peptide #783
10	365	14.1	70	22 ABB33307	Peptide #813 encod

11	365	14.1	70	22	ABB18767	Protein #766 encod
12	365	14.1	70	22	AAW54096	Human brain expres
13	365	14.1	70	22	AAW66486	Human bone marrow
14	365	14.1	70	22	AAW14359	Peptide #793 encod
15	365	14.1	70	22	AAW26772	Peptide #809 encod
16	365	14.1	70	22	AAW02088	Peptide #770 encod
17	365	14.1	70	23	ABG36138	Human peptide enco
18	261	10.1	51	22	ABB28134	Human peptide #785
19	261	10.1	51	22	ABB33309	Peptide #815 encod
20	261	10.1	51	22	ABB18769	Protein #768 encod
21	261	10.1	51	22	AAW54098	Human brain expres
22	261	10.1	51	22	AAW66488	Human bone marrow
23	261	10.1	51	22	AAW14361	Peptide #795 encod
24	261	10.1	51	22	AAW26774	Peptide #811 encod
25	261	10.1	51	22	AAW02090	Peptide #772 encod
26	261	10.1	51	23	ABG36140	Human peptide enco
27	202	7.8	384	21	AAW21784	Arabidopsis thalia
28	202	7.8	384	21	AAW28394	Arabidopsis thalia
29	202	7.8	384	21	AAW49572	Arabidopsis thalia
30	202	7.8	493	21	AAW21783	Arabidopsis thalia
31	202	7.8	493	21	AAW28393	Arabidopsis thalia
32	202	7.8	493	21	AAW49571	Arabidopsis thalia
33	202	7.8	493	23	ABW92337	Herbicidally activ
34	200.5	7.8	360	21	AAW21785	Arabidopsis thalia
35	200.5	7.8	360	21	AAW28395	Arabidopsis thalia
36	200.5	7.8	360	21	AAW49573	Arabidopsis thalia
37	198	7.7	290	21	AAW54540	Zea mays protein f
38	194.5	7.5	282	23	ABW91247	Herbicidally activ
39	192.5	7.5	320	21	AAW43889	Arabidopsis thalia
40	192.5	7.5	327	21	AAW43888	Arabidopsis thalia
41	192.5	7.5	355	21	AAW43887	Arabidopsis thalia
42	191	7.4	816	23	ABW92779	Herbicidally activ
43	190.5	7.4	395	23	ABW93290	Herbicidally activ
44	189.5	7.3	294	23	ABW91093	Herbicidally activ
45	188.5	7.3	322	21	AAW43239	Arabidopsis thalia

ALIGNMENTS

RESULT 1
AAW26706
ID AAW26706 standard; Protein; 504 AA.
XX AC AAW26706;
XX AC
XX DT 14-APR-1998 (first entry)
XX DE Human TAB1 (TAK1 binding protein).
XX KW TAB1; TAK1 binding protein; transforming growth factor-beta;
XX signal transduction; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Misc-difference 52
XX /note= "variant has Arg as residue 52"
XX PN EP803571-A2.
XX PD 29-OCT-1997.
XX PF 24-APR-1997; 97EP-0302808.
XX PR 20-NOV-1996; 96US-0752891.
XX PR 24-APR-1996; 96JJP-0126282.
XX PR 28-OCT-1996; 96JJP-0300856.
XX (UENO/) UENO N.
XX PA Matsumoto K, Nishida E;
XX PI
XX

DR WPI: 1997-515318/48.
 DR N-PSDB; AAT91175.
 PT DNA encoding TAK1 binding protein TAB1 - member of transforming
 PT growth factor beta receptor signal production pathway, which
 PT activates TAK-1 kinase activity upon binding
 XX
 XX
 PS Example 5; Page 17-19; 30pp; English.
 CC This protein comprises human TAB1, a novel member of the
 CC transforming growth factor-beta receptor signal transduction
 CC pathway, which activates TAK-1 kinase activity upon binding. Its
 CC amino acid sequence was deduced from a cDNA clone (see AAT91175)
 CC obtained from a kidney library; a variant TAB1 (see AAM26707) has
 CC Arg rather than Ser at amino acid position 52. Also claimed are:
 CC isolated DNA encoding a protein modified by a substitution,
 CC deletion and/or addition of 1 or more amino acids of the 504-residue
 CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
 CC nucleic acid sequence; (3) isolated DNA encoding a protein
 CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
 CC (4) DNA encoding a fusion protein comprising an above protein or
 CC polypeptide; (5) expression vector comprising an above DNA; and
 CC (6) host cell, preferably a mammalian or yeast cell, transformed by
 CC the expression vector. Cells expressing TAB1 and TAK1 can be used
 CC to screen for TGF-beta signalling pathway inhibitors by contacting
 CC the cells with a test compound, and measuring the TAK1 kinase
 CC activity.
 XX
 XX
 SQ Sequence 504 AA:
 Query Match 100.0%; Score 2580; DB 18; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1e-212;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID AAY09541 standard; Protein; 504 AA.
 XX
 AC AAY09541.
 XX
 DT 21-JUL-1999 (first entry)
 XX
 XX Human TAB1 protein.
 DE
 XX
 KW Human; TAB1; screening; inhibition; TGF-beta;
 KW transforming growth factor beta.
 XX
 XX Homo sapiens.
 OS
 XX
 PN W09921010-A1.
 XX
 PD 29-APR-1999.
 XX
 PF 22-OCT-1998; 98WO-JP04796.
 XX
 PR 22-OCT-1997; 97JP-0290188.
 XX
 PA (CHUS) CHUGAI SEIYAKU KK.
 XX
 PI Ohtomo T, Ono K, Tsuchiya M;
 XX
 DR WPI: 1999-312645/26.
 DR N-PSDB; AAX56278.
 XX
 PT Screening for TGF-beta inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder
 XX
 PS Claim 3; Page 147-149; 195pp; Japanese.
 XX
 CC A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
 CC indications e.g. as TGF-beta signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or cell proliferation prevention inhibitors or
 CC activators, or monocytic migration inhibitors or activators, or
 CC physiological activity induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid beta protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence represents human TAB1.
 XX
 XX
 SQ Sequence 504 AA:
 Query Match 100.0%; Score 2580; DB 20; Length 504;
 Best Local Similarity 100.0%; Pred. No. 1e-212;
 Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 RRIGDYKVKYGYTDIDLSAAKSPKIIAEPEIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Qy 301 GPGQANQIEIAMIDTEFAKQTSLDAAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEIAMIDTEFAKQTSLDAAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGRVVPVSPYSSAQSTKTSVTLSLVMPSQGQWVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVVPVSPYSSAQSTKTSVTLSLVMPSQGQWVNG 420
Qy 421 AHSASTLDEATPTLTNOSPRTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPRTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
RESULT 3
AAV91001
ID AAV91001 standard; Protein; 504 AA.
XX
AC AAV91001;
XX
DT 04-SEP-2000 (first entry)
XX Human TAB-1 protein sequence SEQ ID NO:4.
XX
XX Human, TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression.
XX
OS Homo sapiens.
XX
XX W0200023610-A1.
XX
XX 27-APR-2000.
XX
XX 21-OCT-1999; 99WO-JP05817.
XX
XX 21-OCT-1998; 98JP-0299962.
XX
XX (CHUS) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M, Ontomo T, Sugamata Y, Matsumoto K;
XX WPI; 2000-339707/29.
XX N-PSDB; AAA39106.
XX
XX Method for screening inhibitors of TAK1 signal transduction for
PT suppression of inflammatory cytokine production and use as
PT antiinflammatory agents -
XX
XX Disclosure; Page 90-94; 100pp; Japanese.
XX
XX The present invention describes a method for screening compounds for
CC inhibition of inflammatory cytokine signal transduction by contacting
CC the sample with TAK1 and its receptor TAB1 and selecting for inhibition
CC of TAK1/TAB1 binding. Also described is a method for screening compounds
CC for inhibition of inflammatory cytokine signal transduction in which the
CC inhibition of TAK1 phosphorylation is selected for; and drug
CC compositions for the treatment of inflammatory disorders containing as
CC active component an inflammatory cytokine signal transduction inhibitor.
CC TAK1 is an essential component of the signalling process which results
CC in release of inflammatory cytokines such as interleukin-1 (IL-1),
CC IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
CC for the selection of effective antiinflammatory agents. The present
CC sequence represents human TAB-1, which is used in the exemplification of
XX the present invention.
XX
XX Sequence 504 AA;

Query Match 100.0%; Score 2580; DB 21; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MAQRRLSLQSQPSTDDPLCHLSGVGSASNRYSADKGKTESHPEDSWLKFRSEN 60
Db 1 MAQRRLSLQSQPSTDDPLCHLSGVGSASNRYSADKGKTESHPEDSWLKFRSEN 60
Qy 61 NCFLYGVNGYDGNVTNFAQRLSABELLQGLNAEHAADVRRVLLQAFDVVERSFLS 120
Db 61 NCFLYGVNGYDGNVTNFAQRLSABELLQGLNAEHAADVRRVLLQAFDVVERSFLS 120
Qy 121 IDDALEAKASLQSQPEGVPOHLPPOYOKILERLKTLEIREISGGAMAVVLLNNKLYV 180
Db 121 IDDALEAKASLQSQPEGVPOHLPPOYOKILERLKTLEIREISGGAMAVVLLNNKLYV 180
Qy 181 ANVGTNRALLCKSTVDGLQVTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
Qy 241 RRIGDYKVKYGYTDIDLSAAKSKPIIAEPEIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Db 241 RRIGDYKVKYGYTDIDLSAAKSKPIIAEPEIHGAQPLDGVGTGFLVLMSEGLYKALEAAH 300
Qy 301 GPGQANQIEIAMIDTEFAKQTSLDAAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEIAMIDTEFAKQTSLDAAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTL 360
Qy 361 LVRNFGYPLGEMSOPTSPAPAGRVVPVSPYSSAQSTKTSVTLSLVMPSQGQWVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGRVVPVSPYSSAQSTKTSVTLSLVMPSQGQWVNG 420
Qy 421 AHSASTLDEATPTLTNOSPRTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPRTLQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Qy 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEQSVVTAP 504
RESULT 4
AAV59450
ID AAV59450 standard; Protein; 504 AA.
XX
AC AAV59450;
XX
DT 24-MAR-2000 (first entry)
XX
XX Human TAB1 protein sequence.
XX
XX Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;
KW transforming growth factor-beta activated kinase 1; monocyte migration;
KW TAK1 binding protein 1; extracellular matrix protein production;
KW cell growth inhibitor; beta-amyloid protein deposition;
KW immunosuppression; Transforming growth factor-beta.
XX
XX Homo sapiens.
XX JP11326328-A.
XX
XX 26-NOV-1999.
XX
XX 13-MAY-1998; 98JP-0130378.
XX
XX 13-MAY-1998; 98JP-0130378.
XX (MATS/) MATSUMOTO K.
XX WPI; 2000-078337/07.
XX N-PSDB; AAZ48861.
XX

PT Screening of apoptosis which inhibits combination of the X-linked
PT inhibitor of apoptosis protein -
XX Claim 2; Page 25-26; 43pp; Japanese.
XX
XX This sequence represents the human TAB1 protein.
CC The invention relates to a method for screening a substance inhibiting
CC the formation of a complex between XIAP and TAB1, in which X-linked
CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be
CC tested are contacted with each other and then the presence or formation
CC of a complex between XIAP and TAB1 is detected. The substance can be used
CC as a drug for extracellular matrix protein production enhancement, cell
CC growth inhibition, monocyte migration, physiologically active substance
CC induction, immunosuppression, and beta-amyloid protein deposition. A
CC substance inhibiting the formation of a complex between TAB1 and XIAP as
CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type
CC I and/or type II receptor is useful as a drug.
XX
XX Sequence 504 AA:
SQ
Query Match 100.0%; Score 2580; DB 21; Length 504;
Best Local Similarity 100.0%; Pred. No. 1e-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
DB 61 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
QY 121 IDALAEEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAAVAVLLNNKLYV 180
DB 121 IDALAEEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAAVAVLLNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVTLNVDHTEENDELFRLSQLGDAGKIKOVGIIICQEST 240
DB 181 ANVTNRALLCKSTVDGLQVTLNVDHTEENDELFRLSQLGDAGKIKOVGIIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPHIGAQLDVGTFGLVMSGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPHIGAQLDVGTFGLVMSGLYKALEAAH 300
QY 301 GPQANOEIAAMIDTEFAKOTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEMDTL 360
DB 301 GPQANOEIAAMIDTEFAKOTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEMDTL 360
QY 361 LVNFGYPLGEMSOPTPSAPAAGRVYPVSVSSAQSTSKTSVTLSLVMPSQGMVNG 420
DB 361 LVNFGYPLGEMSOPTPSAPAAGRVYPVSVSVSSAQSTSKTSVTLSLVMPSQGMVNG 420
QY 421 AHAASLTIDETPTLTNOSPFLTLOSTNTHQSSSSDGLFPRPAHSLPGEDEGRVAP 480
DB 421 AHAASLTIDETPTLTNOSPFLTLOSTNTHQSSSSDGLFPRPAHSLPGEDEGRVAP 480
QY 481 YVDFAEFRYRLMSVDHGQSVVTAP 504
DB 481 YVDFAEFRYRLMSVDHGQSVVTAP 504
RESULT 5
AA09550
ID AA09550 standard; protein; 513 AA.
XX
XX AA09550;
AC
XX
XX 21-JUL-1999 (first entry)
DT
XX
XX Human TAB1 protein SEQ ID NO:43.
DE
XX
XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;

KW transforming growth factor beta.
XX
XX Homo sapiens.
OS
XX
XX WO921010-A1.
PN
XX
XX 29-APR-1999.
PD
XX
XX 22-OCT-1998; 98WO-0P04796.
PF
XX
XX 22-OCT-1997; 97JP-0290188.
PR
XX
XX (CHUS) CHUGAI SEIYAKU KK.
PA
XX
XX Ohtomo T, Ono K, Tsuchiya M;
PI
XX
XX WPI; 1999-312645/26.
DR
XX
XX N-PSDB; AAX56310.
PT
XX
XX Example 13; Page 186-188; 195pp; Japanese.
PT
XX
XX A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence represents human TAB1.
XX
XX Sequence 513 AA:
SQ
Query Match 100.0%; Score 2580; DB 20; Length 513;
Best Local Similarity 100.0%; Pred. No. 1e-212;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
DB 1 MAQRSLQSEQPSWTDLPLCHLSGVGSASNSRSYADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
DB 61 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 120
QY 70 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 129
DB 70 NCLYGVFNCGYDGNRVTFVFAQRLSAELLIGQLNAEHAADVRRVLLQAFDVVERSFLES 129
QY 121 IDALAEEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAAVAVLLNNKLYV 180
DB 121 IDALAEEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGGAMAAVAVLLNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVTLNVDHTEENDELFRLSQLGDAGKIKOVGIIICQEST 240
DB 181 ANVTNRALLCKSTVDGLQVTLNVDHTEENDELFRLSQLGDAGKIKOVGIIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPHIGAQLDVGTFGLVMSGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPHIGAQLDVGTFGLVMSGLYKALEAAH 300
QY 301 GPQANOEIAAMIDTEFAKOTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEMDTL 360
DB 301 GPQANOEIAAMIDTEFAKOTSIDAVAQAVDVKRIHSDTFASGGERARFCPRHEMDTL 360
QY 361 LVNFGYPLGEMSOPTPSAPAAGRVYPVSVSSAQSTSKTSVTLSLVMPSQGMVNG 420
DB 361 LVNFGYPLGEMSOPTPSAPAAGRVYPVSVSVSSAQSTSKTSVTLSLVMPSQGMVNG 420
QY 421 AHAASLTIDETPTLTNOSPFLTLOSTNTHQSSSSDGLFPRPAHSLPGEDEGRVAP 480
DB 421 AHAASLTIDETPTLTNOSPFLTLOSTNTHQSSSSDGLFPRPAHSLPGEDEGRVAP 480
QY 481 YVDFAEFRYRLMSVDHGQSVVTAP 504
DB 481 YVDFAEFRYRLMSVDHGQSVVTAP 504
QY 504 YVDFAEFRYRLMSVDHGQSVVTAP 504
DB 504 YVDFAEFRYRLMSVDHGQSVVTAP 504

PS Example 5; Page 19-21; 30pp; English.
XX This protein comprises human TAB1, a novel member of the
CC transforming growth factor-beta receptor signal transduction
CC pathway, which activates TAK-1 kinase activity upon binding. Its
CC amino acid sequence was deduced from a cDNA clone (see AAT91176)
CC obtained from a kidney library; a variant TAB1 (see AAW26706) has
CC ser rather than Arg at amino acid position 52. Also claimed are:
CC isolated DNA encoding a protein modified by a substitution,
CC deletion and/or addition of 1 or more amino acids of the 504-residue
CC TAB1 sequence; (2) DNA which can hybridise with the 1560 bp TAB1
CC nucleic acid sequence; (3) isolated DNA encoding a protein
CC comprising amino acids 21-579 or 437-504 of the 504 TAB1 sequence;
CC (4) DNA encoding a fusion protein comprising an above protein or
CC polypeptide; (5) expression vector comprising an above DNA; and
CC (6) host cell, preferably a mammalian or yeast cell, transformed by
CC the expression vector. Cells expressing TAB1 and TAK1 can be used
CC to screen for TGF-beta signalling pathway inhibitors by contacting
CC the cells with a test compound, and measuring the TAK1 kinase
CC activity.
CC
CC
CC
CC
SQ Sequence 504 AA;

Query Match 99.8%; Score 2575; DB 18; Length 504;
Best Local Similarity 99.8%; Pred. No. 2,7e-212;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 MAQRSSLSQEQPMTDDLPLCHLSGVGSASNSYSADGKTESHPEDSMLFRSEN 60
DB 1 MAQRSSLSQEQPMTDDLPLCHLSGVGSASNSYSADGKTESHPEDSMLFRSEN 60
OY 61 NCFLYGVNGYDGNRVTNFVAORLSAELLGQLNAEHAEDVRVLLQAFDVERSFLBS 120
DB 61 NCFLYGVNGYDGNRVTNFVAORLSAELLGQLNAEHAEDVRVLLQAFDVERSFLBS 120
OY 121 IDALAEKASLSQSPREGVPOHOLPROYOKILRLKTEREISGGAMAVAVALLNNKLYV 180
DB 121 IDALAEKASLSQSPREGVPOHOLPROYOKILRLKTEREISGGAMAVAVALLNNKLYV 180
OY 181 ANVTNRALLCKSTVDGLVOTQNLNVDHTTENEDLFRLSQLGADGKIKQVGIIGQEST 240
DB 181 ANVTNRALLCKSTVDGLVOTQNLNVDHTTENEDLFRLSQLGADGKIKQVGIIGQEST 240
OY 241 RRIQGVKVKYGYTIDLLSAASKPIIAPEPIHGAQPLDGVTFVLWSEGLYKALEAAH 300
DB 241 RRIQGVKVKYGYTIDLLSAASKPIIAPEPIHGAQPLDGVTFVLWSEGLYKALEAAH 300
OY 301 GPGQANQETIAAMIDTEFAKOTSLDAVAQAIVDRVKRIHSDTFASGGERARPCRHEDMTL 360
DB 301 GPGQANQETIAAMIDTEFAKOTSLDAVAQAIVDRVKRIHSDTFASGGERARPCRHEDMTL 360
OY 361 LVNFGPIPGEMSQTPSPAPAGRVYPSVYSAQSTSKTSVTLSLVMSQCGQWNG 420
DB 361 LVNFGPIPGEMSQTPSPAPAGRVYPSVYSAQSTSKTSVTLSLVMSQCGQWNG 420
OY 421 AHSASTILDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 480
DB 421 AHSASTILDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 480
OY 481 YVDFAEFYRLMSVDHGQSIVVTAP 504
DB 481 YVDFAEFYRLMSVDHGQSIVVTAP 504

RESULT 8
ID AAB56692 standard; Protein; 84 AA.
AC AAB56692;
XX
XX
DT 13-MAR-2001 (first entry)
XX
DE Human prostate cancer antigen protein sequence SEQ ID NO:1270.

XX Human; prostate cancer; prostate cancer antigen; detection; diagnosis;
KW neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;
KW vulnary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;
KW antidiabetic; gene therapy; neural; immune; reproductive; renal;
KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder;
KW wound; infectious disease.
XX
OS Homo sapiens.
PN
PN WO200055174-A1.
XX
PD 21-SEP-2000.
XX
XX 08-MAR-2000; 2000MO-US05988.
XX
XX 12-MAR-1999; 99US-0124270.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM;
XX
XX WPI: 2000-587513/55.
DR N-PSDB; AAF15895.
XX
PT Prostate cancer associated gene sequences, referred to as prostate
PT cancer antigens, useful for treatment, prevention, and diagnosis of
PT disorders such as prostate cancer -
XX
PS Claim 11; Page 1693; 2338pp; English.

XX AAF15566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytosolic,
CC cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal,
CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF1514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention.
CC
CC
CC
SQ Sequence 84 AA;

Query Match 15.8%; Score 409; DB 21; Length 84;
Best Local Similarity 97.5%; Pred. No. 1,9e-27;
Matches 78; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 425 STLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 484
DB 5 ATLDEATPTLTNQSPTLTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPGEDEGRVDP 64
OY 485 AEFYRLMSVDHGQSIVVTAP 504
DB 65 AEFYRLMSVDHGQSIVVTAP 84

RESULT 9
ID ABB28132 standard; Peptide; 70 AA.
AC ABB28132;
XX
XX
DT 01-FEB-2002 (first entry)
XX
DE Human peptide #783 encoded by breast cell single exon nucleic acid probe.
XX
XX Human; microarray; single exon probe; gene expression; breast;

```

KW disease; cancer.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
PS Claim 27; SEQ ID NO 11100; 327pp + sequence listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting
CC the probes with a collection of detectably labelled nucleic acids
CC derived from mRNA of human breast, and then measuring the label
CC bound to each probe of the microarray. The probes are useful for
CC verifying the expression of regions of genomic DNA predicted to
CC encode proteins. They are useful for gene discovery, and for
CC determining predisposition and/or prognosing breast disease. Gene
CC expression analysis is useful for assessing the toxicity of chemical
CC agents on cells. The microarray of this invention presents a far greater
CC diversity of probes for measuring gene expression, with far less bias
CC than expressed sequence tag microarrays. The method is suitable for
CC rapid production of functional information from genomic sequence. The
CC present sequence is a peptide encoded by a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 70 AA;

Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKQTSLDAVAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDAVAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 10
ABB33307
ID ABB33307 standard; Peptide; 70 AA.
XX
AC ABB33307;
XX
DT 04-FEB-2002 (first entry)
XX

Peptide #813 encoded by human foetal liver single exon probe.
Human; foetal liver; gene expression; single exon nucleic acid probe.
Homo sapiens.
WO200157277-A2.
09-AUG-2001.
30-JAN-2001; 2001WO-US00669.
04-FEB-2000; 2000US-0180312.
26-MAY-2000; 2000US-0207456.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-0234687.
27-SEP-2000; 2000US-0236359.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2001-483447/52.
Human genome-derived single exon nucleic acid probes useful for
analyzing gene expression in human fetal liver -
Claim 27; SEQ ID NO 25942; 639pp + sequence listing; English.
The invention relates to a single exon nucleic acid probe for
measuring human gene expression in a sample derived from human foetal
liver. The single exon nucleic acid probes may be used for predicting,
measuring and displaying gene expression in samples derived from human
fetal liver. The present sequence is a peptide encoded by a single exon
nucleic acid probe of the invention.
Note: The sequence data for this patent did not form part of the
printed specification, but was obtained in electronic format directly
from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ Sequence 70 AA;

Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEFAKQTSLDAVAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDAVAQAVVDVRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

QY 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 11
ABB18767
ID ABB18767 standard; Protein; 70 AA.
XX
AC ABB18767;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #766 encoded by probe for measuring heart cell gene expression.
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.

```

```
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00666.
PF
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
PT
XX Claim 15; SEQ ID No 20537; 530bp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 70 AA;
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTPFASGGERARFCPRHEDMTLLVNFQYPLGE 371
DB 1 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTPFASGGERARFCPRHEDMTLLVNFQYPLGE 60
QY 372 MSQPTPSPAP 381
DB 61 MSQPTPSPAP 70
RESULT 12
AAM54096
ID AAM54096 standard; Protein; 70 AA.
XX
XX AAM54096;
AC
XX 05-NOV-2001 (first entry)
DT
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 26201.
DE
XX Human; brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
OS
XX WO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00667.
PF
```

```
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
PT
XX Example 4; SEQ ID NO: 26201; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 70 AA;
Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 312 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTPFASGGERARFCPRHEDMTLLVNFQYPLGE 371
DB 1 MIDTEFAKQTSLDVAQAQAVVDRVKRIHSDTPFASGGERARFCPRHEDMTLLVNFQYPLGE 60
QY 372 MSQPTPSPAP 381
DB 61 MSQPTPSPAP 70
RESULT 13
AAM66486
ID AAM66486 standard; Protein; 70 AA.
XX
XX AAM66486;
AC
XX 06-NOV-2001 (first entry)
DT
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 26792.
DE
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
OS
XX WO200157276-A2.
PN
XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00668.
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
```

XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 26792; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.
 XX
 SQ Sequence 70 AA;
 Query Match 14.1%; Score 365; DB 22; Length 70;
 Best Local Similarity 100.0%; Pred. No. 8.5e-24;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 312 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
 Db 1 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60
 QY 372 MSQPTSPAP 381
 Db 61 MSQPTSPAP 70
 RESULT 14
 AAM14359
 ID AAM14359 standard; Protein; 70 AA.
 AC AAM14359;
 XX 12-OCT-2001 (first entry)
 DE Peptide #793 encoded by probe for measuring cervical gene expression.
 XX Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX Homo sapiens.
 OS
 PN WO200157278-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US006670.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human cervical epithelial cells -
 XX
 PS Claim 27; SEQ ID No 19185; 487pp; English.
 CC The present invention relates to human single exon nucleic acid probes

CC (SENP: see AA110068-AA128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 70 AA;
 Query Match 14.1%; Score 365; DB 22; Length 70;
 Best Local Similarity 100.0%; Pred. No. 8.5e-24;
 Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 312 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
 Db 1 MIDTEFAKQTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60
 QY 372 MSQPTSPAP 381
 Db 61 MSQPTSPAP 70
 RESULT 15
 AAM26772
 ID AAM26772 standard; Protein; 70 AA.
 XX
 AC AAM26772;
 XX 17-OCT-2001 (first entry)
 DE Peptide #809 encoded by probe for measuring placental gene expression.
 XX Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder.
 XX Homo sapiens.
 OS
 PN WO200157272-A2.
 XX 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US006663.
 XX 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 27; SEQ ID No 27041; 654pp; English.
 CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see AA131315-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.
 XX

SQ Sequence 70 AA;

Query Match 14.1%; Score 365; DB 22; Length 70;
Best Local Similarity 100.0%; Pred. No. 8.5e-24;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 312 MIDTEPAKQTSIDAVAQAVVDVKRIHSPTFASGGERARFCPRHEDMTLLVNFQYPLGE 371
|||
DB 1 MIDTEPAKQTSIDAVAQAVVDVKRIHSPTFASGGERARFCPRHEDMTLLVNFQYPLGE 60
QY 372 MSOPTPSPAP 381
|||
DB 61 MSOPTPSPAP 70

Search completed: December 9, 2002, 22:47:18
Job time : 55.0526 secs

QY 9 LQSEQQSWTDDLPCHLSGVGSASNRYS---ADKGTEGSHPPEDSWLKFRSENN--CF 63
Db 8 LYSNQPO-TVEAP---ASGGGLSQNGKFSYGYASSAGKRSMEDFFETRIDGINGEIVG 63
QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHA--ADVRVLLQAFDVVERSFLESI 121
Db 64 LFGVFDGHHGARAAYVVRHLFSNLI-----THPKFISDTKSAITDAYNHTSELKSE 117
QY 122 DDLAEKASLOSQLPQVPOHQLPOYQKILERLKTLEISGAMAVAVLLNNKLYVA 181
Db 118 N-----SHNRDAGSTASTAILVGDRLVVA 141
QY 182 NVGTRALLCKSTVDGLQVTLQNLVDHTTENEDELFRLSQLDAGKIKQVGIICGOSTR 241
Db 142 NVGDSRAVISR---GGKAIASVRDHKPDQSDERERENAG---GFVWAGVLA---VSR 191
QY 242 RIGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVGTFLVLMSEGLYKALENAHG 301
Db 192 AFGORLLK-----QYVADPEIQ-BEKIDDTLEFLILASDGLWDVF----- 231
QY 302 PQOANQEIAMI---DTEFAKQTSLDAAVAVVDRVKRIHSDTFASGGERARFCPRHED 357
Db 232 ---SNEAAVAMVKEVDEP-----DSAKLVGEAIKRGSA-----N 265
QY 358 MTLVVRNPGYPLGEMSOPTSPAPAGRRVVPVSVSSAOSTSKTSVTLSLVNPSQGM 417
Db 266 ITCVVVRP-----LEKGSASSSHISSSSSSKEAKEMPPLGDL 301
QY 418 VNGAHSASTLD---EATPTLTNOSPILTLQSTNTHTQSSSSSD 458
Db 302 AISSNEAKQVOIGNGKPNVTRKPDATASRSTDTLTLERNSTVD 346
RESULT 5
T05680
hypothetical protein F20M13.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 20-Jun-2000
C:Accession: T05680
R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15420
A:Accession: T05680
A:Molecule type: DNA
A:Residues: 1-395 <BEV>
A:Cross-references: EMBL:AL035540
A:Experimental source: cultivar Columbia, BAC clone F20M13
C:Genetics:
A:Map position: 4
A:Introns: 102/1; 224/3; 303/3
A:Note: F20M13.80
C:Superfamily: Arabidopsis thaliana hypothetical protein F20M13.80

Query Match 7.4%; Score 190.5; DB 2; Length 395;
Best Local Similarity 24.9%; Pred. No. 7.6e-06;
Matches 104; Conservative 60; Mismatches 150; Indels 111; Gaps 20;
QY 4 QRRSLLOEQQPSWTDLDPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSENNCF 63
Db 49 QANSLLEDQSQ-----LESGLSHSDSGPFT----- 75

QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHAADVRVLLQAFDVVERSFLESI 123
Db 76 FVGVDYDGHGGPETSFRFINDHMFPH--LKRTAQ-QCMSSEVIKAKFAQTEGFLSIVTN 132
QY 124 ALAEKASLOSQLPQVPOHQLPOYQKILERLKTLEISGAMAVAVLLNNKLYVANV 183
Db 133 -----QFOTRPOIATV-----GSCCLVSVICDGLKYVANA 162
QY 184 GTNRALLCK-STVDG-LQVTLQNLVDHTTENEDELFRLSQLDAGKI-----KQVGI 233
Db 163 GDSRAVLQGVWRVTGEHAATQLSAEHNASIESVRRELQALHPDPIVLKHNWVRVKGI 222

QY 234 ICQESTRRIGDYKVKYG-----YTDIDLLSAASKPII-AEPEI--HGAQPLDGVGT 283
Db 223 I---QVSRIGDYLKRFSEFNREPIYAKFRLSPF-SRPLLSAEPAITVHTLPHD---Q 275
QY 284 FLVLMSEGLYKALEAAGHGPQANQEIAMIDTE----FAKQTSLDAAVAVVDRVKRIHS 339
Db 276 FIICASDGLWEHM-----SNQEAVIDVQNHPRNGIAKRLVKVALQEAACKREMYSD 327
QY 340 DTFASGGERARFCPRHEDMTLLVRNPGYPLGEMSOPTSPAPAGRRVVPVSVSS-AQ 398
Db 328 LKKIDRGVRRHF---HDDITIVVFFDTNL--VSRGSMRLRGPAVSVRGAGVNLPHNTLAP 382
QY 399 STSKT 403
Db 383 CITPT 387

RESULT 6

A55804
phosphoprotein phosphatase (EC 3.1.3.16) 2c, membrane-bound - Paramesium tetraurelia
C:Species: Paramesium tetraurelia
C:Date: 23-Mar-1995 #sequence_revision 05-Apr-1995 #text_change 07-Dec-1999
C:Accession: A55804
R:Klumpp, S.; Hanke, C.; Donella-Deana, A.; Beyer, A.; Kellner, R.; Pinna, L.A.; Schultz, J. Biol. Chem. 269, 32774-32780, 1994
A:Title: A membrane-bound protein phosphatase type 2C from Paramesium tetraurelia. Purified from Paramesium tetraurelia.
A:Reference number: A55804; MUID:95105156; PMID:7806499
A:Accession: A55804
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-300 <KLU>
A:Cross-references: GB:Z36985; NID:g537421; PID:el192609; PID:g2654382
C:Genetics:
A:Genetic code: SGC5
C:Keywords: phosphoric monoester hydrolase

Query Match 7.4%; Score 190; DB 2; Length 300;
Best Local Similarity 24.5%; Pred. No. 5.3e-06;
Matches 66; Conservative 51; Mismatches 104; Indels 48; Gaps 9;
QY 64 LYGVFNQYDGNRVNFAQRLSAELLQLQNAEHAADVRVLLQAFDVVERSFLESI 123
Db 52 VFGVFDGHHGREGVAQFVEKHFDVDELLKNK-----NFKQKFEALKE 93
QY 124 ALAEKASLOSQLPQVPOHQLPOYQKILERLKTLEIS--GGAMAVAVLLNNKLYVAN 182
Db 94 TPLKMDL-----LLTPGQKELNQYKATDTDESAGCTANVALIYKNTLYVAN 142
QY 183 VGTNRALLCKSTVDGLQVTLQNLVDHTTENEDELFRLSQLDAGKIKQVGIICGQESTRR 242
Db 143 AGDSRSLCRNTN---HDSVVDHKKPDNPEKSRIERAG---GFVSDGRVNGNINLSRA 195

QY 243 IGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQ--PLDGVTFGLVLMSEGLYKALEAAH 300
Db 196 LGDLEYKR-----DNKLRNQELIILALPDVKKTELTPQD---KFLMGCDGVFTLNHOE 247
QY 301 GPGQANQEI-AMIDTEFAKQTSLDAAVQ 328
Db 248 LLKQVNSTIGQAQVTELLKKAEDLLDQ 276

RESULT 7

P96355
T16E15.10 protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 27-Nov-2001
C:Accession: P96355
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, Ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, C.J.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: F86355
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-281 <STO>
 A:Cross-references: GB:AE005172; NID:99392666; PIDN:AAF87263.1; GSPDB:GN00141
 C:Genetics:
 A:Map position: 1

Query Match 7.3%; Score 188.5; DB 2; Length 281;
 Best Local Similarity 24.8%; Pred. No. 6.1e-06;
 Matches 83; Conservative 59; Mismatches 88; Indels 105; Gaps 17;

QY 30 GSASNR---YSADG-----KGTSEHPED-----SWLKFSENNCFIYGVFNGYDG 73
 DB 16 GSSSRSGKSGSDSCMIKXGSLVKGKANHPEVDYHNFNIO-DHELGLFAIYDGHMG 74
 QY 74 NRVTFVAGQRLSAELLGQLNAEHAE--ADYRVLLQAFDVERSFLESIDALAEKASL 131
 DB 75 DSVPAVLQKRLPSNLT-----KEGEFWDPRRSIAKAYEKTDAILNSSD----- 120
 QY 132 QSLQEGVPOHQLPPOYQKILERLKTLEISGAMAVAVALLN-NKLYVANVGTNRALL 190
 DB 121 -----LGR-----GSGTAVATILINRGKLIANVGDNRAYL 151
 QY 191 CKSTVDGLVQLNVNDH--TTEN---EDBLFSLQGLDAGKIKVGIICQESRRICD 245
 DB 152 S-----HGAITQMSDHEPRTERRSIEDRGFVSNLPGVPRVN-----GQ 193
 QY 246 YVKKVGYTDIDLSAAKSKPIIAEPEIHGAQPLDVGTFGLVMSSEGLYK-----ALBAA 299
 DB 194 LVAISRAFGDKGLHTLSSPDIKEATV-----DSQTDVLLASDGINKWMNEAMEIA 247
 QY 300 H--GFGANOIEIAMIIDTEFAKQTSLDVAQAAYV 331
 DB 248 RRVKDPQAKAKELTA---EALRRESKDISCVV 278

RESULT 8

T06308
 protein phosphatase 2C homolog F11C18.60 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 15-Jun-2001
 C:Accession: T06308
 R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Daseville, R.; De Clerck, R.; De ewes, H.W.; Mayer, K.F.X.; Schueller, C.
 A:Submitted to the Protein Sequence Database, April 1999
 A:Reference number: Z15589
 A:Accession: T06308
 A:Molecule type: DNA
 A:Residues: 1-357 <BEV>
 A:Cross-references: EMBL:AL049607; GSPDB:GN00062; ATSP:F11C18.60
 A:Experimental source: cultivar Columbia; BAC clone F11C18
 C:Genetics:
 A:Gene: ATSP:F11C18.60
 A:Map position: 4
 A:introns: 39/3; 61/1, 97/2, 148/3, 190/3, 232/1, 257/3, 275/2; 293/3
 C:Superfamily: human phosphoprotein phosphatase 1A

Query Match 7.3%; Score 188.5; DB 2; Length 357;
 Best Local Similarity 24.6%; Pred. No. 9e-06;
 Matches 85; Conservative 65; Mismatches 145; Indels 51; Gaps 14;

QY 59 ENNCPLTVGVNGYDGNRVTFVAGQRLSAELLGQLNAEHAEADVRRVLLQAF---DVV-- 113
 DB 48 DNTSFL-GVYDGHGKIVSKFCAYLHQVLSD---AFAAGDVGTSLOKAFRRMDMMQ 103
 QY 114 -ERSFLESIDALAEKASLQSLPEGV-----PQHQLPPOYQKILERLKTLEISGGA 166

DB 104 GQGRMELA--VLGKINKFSGMTIGLIMSPPSGSANKPDMAFEEBPHDFAPGNSG 161
 QY 167 MAVAVALLNNKLYVANVGTNRALLCKSTVYDGLQTVLNVDTTENDELFRISQLGDAG 226
 DB 162 TVCAVAVRDQQLFVNAVAGSRVCVSRKN---QAVNLSRDHKFDLEAKERI---LKAG 213
 QY 227 KIKVQVILIQG-ESTRRIGDYKVKGYTDIDLSAAKSKPIIAEPEIHGAQPLDVGTFGL 285
 DB 214 GFLHAGRVNGSINLSRALGDMFKQ---NKLPEBKQIVTASPDVNVTELCD-DDFL 267
 QY 286 VLMSEGLYKALBAHGPQANOIEIAMIIDTEFAKQTSLDVAQAAYVDRKIRHSPTFASG 345
 DB 268 VLACDGINCM-----TSQQLVDFIHEQLNSETKLSVCEKTLDCLAANT---SG 315
 QY 346 GERARFCPEHEDMTLVIRNFGYPLGEMSOPTSPAPPAAGRVYPS 391
 DB 316 GEGC-----DNMTWILVRFKNPTSETELKPEASQAEGNHDPESS 355

RESULT 9

S61650
 hypothetical protein YOR090C - yeast (*Saccharomyces cerevisiae*)
 N:Alternate names: hypothetical protein O3157; hypothetical protein YOR3157c
 C:Species: *Saccharomyces cerevisiae*
 C>Date: 09-Mar-1996 #sequence_revision 12-Apr-1996 #text_change 19-Apr-2002
 C:Accession: S61650; S66975
 R:Benes, V.; Andrade, M.A.; Rechmann, S.; Teodoru, C.; Banrevi, A.; Sander, C.; Valencia submitted to the EMBL Data Library, December 1995
 A:Description: Nucleotide sequence and analysis of a 130 kb fragment of yeast chromosome
 A:Reference number: S61643
 A:Accession: S61650
 A:Molecule type: DNA
 A:Residues: 1-572 <BEN>
 A:Cross-references: EMBL:X94335; NID:g1262139; PID:e217726; PID:g1164936
 R:Voss, H.; Benes, V.; Rechmann, S.; Teodoru, C.; Schwager, C.; Paces, V.; Ansoerge, W. submitted to the Protein Sequence Database, July 1996
 A:Reference number: S66965
 A:Accession: S66975
 A:Molecule type: DNA
 A:Residues: 1-572 <VOSS>
 A:Cross-references: EMBL:Z74998; NID:g1420258; PID:e252349; PID:g1420259; MIPS:YOR090C
 A:Experimental source: strain S288C
 C:Genetics:
 A:Gene: SGD:PTCS
 A:Cross-references: SGD:S0005616
 A:Map position: 15R

Query Match 7.1%; Score 183.5; DB 2; Length 572;
 Best Local Similarity 22.9%; Pred. No. 4.3e-05;
 Matches 107; Conservative 74; Mismatches 159; Indels 127; Gaps 21;

QY 46 SHPEDSMWK-----FRSEN-----NCFLYGVNGYDGNRVTFVAGQRLSAELL----- 89
 DB 162 NHPEDDHVEQIITPIESSEDGKSIEKDIYFFGIFGHGQ---PTSEKLSDLVRVYA 217
 QY 90 --LGQL---NAEHAEADVRRVLLQAFDVERSFLESIDALAEKASLQSLQPGVPOHQL 144
 DB 218 YQLQVYDQNKTVFHSPPNQLISA---ISKGLKLDNDLVISFRLRQ----- 264
 QY 145 PPOYQKILERLKTLEISGAMAVAVALLN-NKLYVANVGTNRALLCKSTVDG-LQYT 201
 DB 265 DPNNNTNANTLPAI-----SGSCHLLSLVNSTSLIKVANVTGSRALLIGLDNRGMWTVK 319
 QY 202 QLVNDHTTENDELFRISQLGLDAGKIKVGIICQ-ESTRRIGDYKVKGYTD-----ID 256
 DB 320 SLSTDQGDNDLDEVRRIRKHPGPVIRNGRILGSLQPSRAGDYRKKEVDGKPLSD 379
 QY 257 LLSAA-----SKSP-IIAEPEIHGAQPLDVGTFGLVMSSEGLYKALBAHGPQ 304
 DB 380 LPEVAKULFRREPRDFTPPVVTAPEVITSK-IGENTKVMVSGSLFELL----- 430
 QY 305 ANQEIAMIIDTEFAKQTSL-----DAVAQAAYVDRKIRHSPTFASG 346

C:Genetics:
A:Gene: At2g34740, T29F13.5
A:Map position: 2
A:Introns: 46/3; 139/1; 190/3

Query Match 6.9%; Score 177.5; DB 2; Length 239;
Best Local Similarity 25.9%; Pred. No. 2.9e-05;
Matches 78; Conservative 50; Mismatches 90; Indels 83; Gaps 14;

Qy 39 ADGKTESHPEDSWLKFRENNCFLYGVFNQDGNRVTFVQAQRLSAELLQLNAEHA 98
Db 7 ADTKIVYKGH-----NLGIYAFDGHSGSDVADYQNLNLPNIL-----S 45
Qy 99 EADVRVLLQAFDVERSGLESIDDALEKASLQSQLPBGVQHQLPPOYOKILERLKT 158
Db 46 QPFDVFNPKKA---IKRAY-KSTDYIIL-----QNVVGR----- 76
Qy 159 EREISGGAAVAVVLLN-NKLVA NVGTNRALLCKSTVDCLOVQTOLNHTTENDELFR 217
Db 77 -----GGSTAIVTAIVDGGKIVAVNGDSRALICR---ESDVVKQITVDHEPKERDLVK 128
Qy 218 LSQGLDACKIKQVGIICGO-ESTRIGDYKVKGYTIDILLSAKSPITAEPIHGAQ 276
Db 129 -SKGGSVSKPGQNVPRVDSQLAMTRAFSGDGLKEHISVI-----PNIETAIHND- 177
Qy 277 PLDGVTFGLVLMSEGLYKAL-----EAAHGPQANQETIAMITPEFAKQTSLDVAQAV 330
Db 178 -----TKFILLASDGLMKVMSNDEVWDQIKKRGNAEAAKMLIDKALAR-GSKDIDISCV 231
Qy 331 V 331
Db 232 V 232

RESULT 13
S55457
phosphoprotein phosphatase (EC 3.1.3.16) 2C - Arabidopsis thaliana
N:Alternate names: protein phosphatase 2C
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 01-Aug-1995 #sequence_revision 01-Sep-1995 #text_change 01-Mar-2002
C:Accession: S55457; S53581
R:Yamamoto, M.
submitted to the EMBL Data Library, August 1994
A:Reference number: S55457
A:Accession: S55457
A:Molecule type: mRNA
A:Residues: 1-399 <YAM>
A:Cross-references: EMBL:D38109; NID:G633027; PID:dl007865; PID:G633028
R:Kurumori, T.; Yamamoto, M.
Nucleic Acids Res. 22, 5296-5301, 1994
A:Title: Cloning of cDNAs from Arabidopsis thaliana that encode putative protein phosphatase
A:Reference number: S53581; MUID:95116318; PMID:7816619
A:Accession: S53581
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 82-399 <KUR>
A:Cross-references: GB:D38109
C:Superfamily: Arabidopsis thaliana hypothetical protein F7A7.220
C:Keywords: phosphoric monoester hydrolase

Query Match 6.8%; Score 175; DB 2; Length 399;
Best Local Similarity 25.5%; Pred. No. 9.8e-05;
Matches 93; Conservative 58; Mismatches 117; Indels 96; Gaps 20;
Qy 7 SLLOSQQQSWTDDLPYLCHLSGVGSASNRYSADGKTES--HPEDSWLKFRENNCFLY 65
Db 91 SYTEAE--SFESDVP-----KIGTTSVCGRRDMDVASHIP---SFLQNSRNHH--Y 138
Qy 66 GVFNQDGNRVTFVQAQRL-----SABLLQLNAEHAEDVRVLLQAFDVERSGFLE 119
Db 139 GVFDGHCCHVAEKCRERLHDIVKKEVEVMASD---EMTETWVKSFQKMDKEVSQRECNL 195
Qy 120 STDDAL-AEKASLQSQLPBGVQHQLPPOYOKILERLKTLEKEISGGAAVAVVLLNKL 178

Db 196 VVNGATRSKMNKSCRCIQS-----PCQDA-----GSTAVVSVTPPEKI 234
Qy 179 YVANVTNRALLCKSTVDCLOVQTOLNVDHTTENDELFRLSQGLDACKI-----KQVG 232
Db 235 IVNCGSGRAVLGRNV-----AIPLSVDHKPRPBDLIRIQAG---GAVIWDGARVLG 287
Qy 233 IICQESTRIGYKVKGYTIDILLSAKSPITAEPIHGAQPLDGVTFGLVMSGL 292
Db 288 VLA---MSRAIGDNYLK-----PVVIPDPEVTWTDRTD-EDECLILASDGL 329
Qy 293 YKAL--EAA-----HGGQANQELIA-----MIDTEFAKQTSLDVAQAVVD 332
Db 330 WDVVNETACGAVMCLRGAGGDSDAHNAACSDALLITLALAKSS-DNVSVVVD 388
Qy 333 RVKR 336
Db 389 LRKR 392

RESULT 14
T47644
protein phosphatase 2C-like protein - Arabidopsis thaliana
N:Alternate names: protein T15C9.50
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 21-Jul-2000
C:Accession: T47644
R:Mewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z24470
A:Accession: T47644
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-409 <MEW>
A:Cross-references: EMBL:AL132970
A:Experimental source: cultivar Columbia; BAC clone T15C9
C:Genetics:
A:Map position: 3
A:Introns: 46/1; 336/3
A:Note: T15C9.50
C:Superfamily: Arabidopsis thaliana hypothetical protein F20M13.80

Query Match 6.8%; Score 175; DB 2; Length 409;
Best Local Similarity 25.2%; Pred. No. 0.0001;
Matches 101; Conservative 56; Mismatches 132; Indels 112; Gaps 21;

Qy 25 HLSG-----VGSASNRYSADGKTESHP-----PEDSWLKFRENNCFLYGVFNQD 72
Db 71 HITGEPMAVVOANN--LEEDHSQLESGLSIHESGPEATFV-----GVYDGHG 117
Qy 73 GNRVTNFVAQRLSAELLQLNAEHAEDVRVLLQAFDVERSGFLESIDDALEKASIQ 132
Db 118 GPFAARFVNDRLF-----YNIKRYTSQRGM---SPDVIIRGFVATEEFL----- 160
Qy 133 SQLPBGVQHQLPPOYOKILERLKTLEKEISGGAAVAVVLLNKLKYVANVTNRALLCK 192
Db 161 -----GLVQ-----EQWKTPQIASVGACCLVIGVNGLLYVANADSDRVVLGK 204
Qy 193 --STVDLCLOVQTOLNVDHTTENDELFRLSQGLDACKI-----KQVGIIICQESTR 242
Db 205 VANPFEKKAQVQSTEHNASIBSVREELRLHRDDNIVVLGHKVKRVGII---QVSHS 261
Qy 243 IGD-YKVKGYTIDILLSAK-----SKPIT-AEPEI--HGAQPLDGVTFGLVMSGLY 293
Db 262 IGDVYLKRAEFNGEPRLPKFRVPERFEKPIMAEPTITVHKIHPED---QFLIFASDGLM 318
Qy 294 KALEAHHGPQANQETIAMITD---EFAKQTSLDVAQAVDVRVKRHSDFPASGGERA 349
Db 319 EHL-----SNQEAVIDVNSCPRNGVARKLVKALQAAKREWRYSDELKEIRGIIR 370
Qy 350 RFPCEHEDMTLLV-----RNFGVPLGEMSQPTSPAPAGG 385
Db 371 HF--HDDITVIVVFLHATNFA-----TRTPISTVAGG 399

Search completed: December 9, 2002, 22:59:41
Job time : 31.8532 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:45:29 ; Search time 54.4488 Seconds
(without alignments)
1907.257 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRRLSQEQPQSWTDD.....AEFYRLWSVDHGQSVVTAP 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 21:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mmc:.*
8: sp_organelle:.*
9: sp_phase:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_rvirus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2512	97.4	500	11 Q8ROD1	Q8ROD1 mus musculus
2	1924	74.6	498	13 Q73614	Q73614 xenopus lae
3	275	10.7	52	11 Q9CV62	Q9CV62 mus musculus
4	258.5	10.0	386	5 Q93375	Q93375 caenorhabdi
5	205.5	8.0	311	10 Q8RXV3	Q8RXV3 arabidopsis
6	205	7.9	319	10 Q8S8Z0	Q8S8Z0 mesembryant
7	202	7.8	493	10 Q9LUS8	Q9LUS8 arabidopsis
8	194.5	7.5	282	10 Q9S9Z7	Q9S9Z7 arabidopsis
9	193.5	7.5	377	10 Q9LNF4	Q9LNF4 arabidopsis
10	192.5	7.5	355	10 Q81716	Q81716 arabidopsis
11	191.5	7.4	243	10 Q8S3P1	Q8S3P1 oryza sativ
12	191	7.4	816	10 Q9MLV6	Q9MLV6 arabidopsis
13	190.5	7.4	283	10 Q9AE3	Q9AE3 arabidopsis
14	190.5	7.4	348	10 Q9LEW5	Q9LEW5 arabidopsis
15	190.5	7.4	395	10 Q9SZN2	Q9SZN2 arabidopsis
16	190.5	7.4	400	10 Q8W4N8	Q8W4N8 arabidopsis

17	188.5	7.3	281	10 Q9LME4	Q9LME4 arabidopsis
18	188.5	7.3	357	10 Q9S253	Q9S253 arabidopsis
19	187.5	7.3	379	10 Q9LSN8	Q9LSN8 arabidopsis
20	186.5	7.2	385	10 Q9FKX4	Q9FKX4 arabidopsis
21	186	7.2	378	10 Q8S2S5	Q8S2S5 thellungiel
22	184.5	7.2	282	10 Q9ZPL8	Q9ZPL8 lotus japon
23	184.5	7.2	420	10 Q8VZD9	Q8VZD9 arabidopsis
24	183.5	7.1	572	3 Q12511	Q12511 saccharomyc
25	183	7.1	305	10 Q942N4	Q942N4 oryza sativ
26	182.5	7.1	420	10 Q94AT1	Q94AT1 arabidopsis
27	182	7.1	271	10 Q9SMJ9	Q9SMJ9 sporobolus
28	181.5	7.0	238	10 Q9C9R2	Q9C9R2 arabidopsis
29	181	7.0	307	10 Q9FGM3	Q9FGM3 arabidopsis
30	180	7.0	355	10 Q9M8J1	Q9M8J1 arabidopsis
31	180	7.0	366	10 Q9FOY2	Q9FOY2 zea mays (m
32	178.5	6.9	361	10 Q9LZ09	Q9LZ09 arabidopsis
33	178	6.9	392	4 Q9H0C8	Q9H0C8 homo sapien
34	177.5	6.9	239	10 Q64583	Q64583 arabidopsis
35	175	6.8	384	10 Q94CL8	Q94CL8 arabidopsis
36	175	6.8	409	10 Q9M2W1	Q9M2W1 arabidopsis
37	173.5	6.7	379	10 Q9SDI2	Q9SDI2 arabidopsis
38	172.5	6.7	528	10 Q93YS2	Q93YS2 arabidopsis
39	172	6.7	464	10 Q9FXE4	Q9FXE4 arabidopsis
40	169.5	6.6	392	4 Q96NT4	Q96NT4 homo sapien
41	169.5	6.6	392	11 Q9Z1Z6	Q9Z1Z6 rattus norv
42	169.5	6.6	416	10 Q9FEW0	Q9FEW0 nicotiana t
43	169	6.6	389	10 Q81773	Q81773 arabidopsis
44	169	6.6	416	10 Q9FLI3	Q9FLI3 arabidopsis
45	168	6.5	423	10 Q9MLV8	Q9MLV8 arabidopsis

ALIGNMENTS

RESULT 1

Q8ROD1
ID Q8ROD1 PRELIMINARY; PRT; 500 AA.
AC Q8ROD1;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Similar to mitogen-activated protein kinase kinase 7
DE interacting protein 1 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-KIDNEY;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC027054; AAH27054.1; -.
KW Kinase.
FT NON TER 1
SQ SEQUENCE 500 AA; 54413 MW; 97529D2E2AE4DA34 CRC64;

Query Match 97.4%; Score 2512; DB 1; Length 500;
Best Local Similarity 97.2%; Pred. No. 4.2e-168;
Matches 488; Conservative 9; Mismatches 3; Indels 2; Gaps 1;

QY 3 AQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSNLKFSENNC 62

Db 1 AQRRLSQEQPQSWTDDLPLCHLSGVGSASNRYSADGKGTESHPPEDSNLKFSENNC 60

QY 63 FLYGVFNGYDGNRVTFNFAVQRLSAELLQLQNAEHAEDVRRVLQAQDVVVERSFLSID 122

Db 61 FLYGVFNGYDGNRVTFNFAVQRLSAELLQLQNAEHAEDVRRVLQAQDVVVERSFLSID 120

QY 123 DALAKASLQSQLPQGVQHQLPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYVAN 182

Db 121 DALAKASLQSQLPQGVQHQLPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYVAN 180

QY	183	VGNTRALLCKSTVDLQVLTQNLNVDHTTENEDELFTLSQLGDAGKIKQVGIIICQESTRR	242
Db	181	VGNTRALLCKSTVDLQVLTQNLNVDHTTENEDELFTLSQLGDAGKIKQVGIIICQESTRR	240
QY	243	IGDVKKVGYVDIDLLSAASKSPIIAEBEIHGAQPLDGVTFVLMSBGLYKALEAHP	302
Db	241	IGDVKKVGYVDIDLLSAASKSPIIAEBEIHGAQPLDGVTFVLMSBGLYKALEAHP	300
QY	303	GOANDEIAMIIDTEFAKQTSLDVAQAVVDKRIHSDTFAGGERARPCPHNEMTLV	362
Db	301	GOANDEIAMIIDTEFAKQTSLDVAQAVVDKRIHSDTFAGGERARPCPHNEMTLV	360
QY	363	RNFQVPLAEMISQPTSPAPAGGRVYPVPSYSAQSTSKTSVTLSLWPSQGMVNGAH	422
Db	361	RNFQVPLAEMISQPTSPAPAGGRVYPVPSYSAQSTSKTSVTLSLWPSQGMVNGAH	418
QY	423	SASTLDEATPTLTNOSPPTLTLOSTINTHTOSSSSSDGGLFRSPAHSLPPEGDEGVEPV	482
Db	419	SASTLDEATPTLTNOSPPTLTLOSTINTHTOSSSSSDGGLFRSPAHSLPPEGDEGVEPV	478
QY	483	DPAEFYRLMSVDHGEQSVVTAP 504	
Db	479	DPAEFYRLMSVDHGEQSVVTAP 500	
RESULT 2			
ID	073614	PRELIMINARY; PRT; 498 AA.	
AC	073614;		
DT	01-AUG-1998 (TREMBlrel. 07, Created)		
DT	01-AUG-1998 (TREMBlrel. 07, Last sequence update)		
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)		
DE	Tabl.		
OS	Xenopus laevis (African Clawed Frog).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;		
OC	Xenopodidae; Xenopus.		
CC	NCBI_TaxID=8355;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RX	MEDLINE=98130593; PubMed=9463380;		
RA	Shibuya H., Iwata H., Masuyama N., Gotoh Y., Yamaguchi K., Irie K.,		
RA	Matsunoto K., Nishida E., Ueno N.;		
RT	"Role of TAK1 and TAB1 in BMP signaling in early Xenopus		
RT	development."		
RL	EMBO J. 17:1019-1028(1998).		
DR	EMBL; U92031; AAC14009.1; -		
DR	InterPro; IPR001932; PP2C-like.		
DR	Pfam; PF00481; PP2C: 1.		
DR	SMART; SM00332; PP2Cc; 1.		
DR	SEQUENCE 498 AA; 54450 MW; 654ABBD2535324 CRC64;		
Query Match 74.6%; Score 1924; DB 13; Length 498;			
Best Local Similarity 75.0%; Pred. No. 8e-127;			
Matches 378; Conservative 54; Mismatches 64; Indels 8; Gaps			
QY	1	MAAQRSLIQSFQPSWTDLPLCHLSGYGSANRSYSADGKTESHPPEDSWLKRSEN	60
Db	1	MAAPRRNLHSQ---SWTDPLPLCNISGVGSANQTYNSBGLCKDHPYEDNWKIRGDN	57
QY	61	NCFLVGVNFGYGNRTNVAORLSALLGOLNAEHAEDVARYVLOAFDVERSEFLS	120
Db	58	NYLVGVNFGYGNRTNVAORLSALLGOLDPVTVDAEHKVLLOAFDVERSEFLS	117
QY	121	IDDALEAKSLQSQPEGVPOHQLPQYOYLERLTLERREISGAMAVAVALLNNKLVY	180
Db	118	IDDCLEAKSLSLQPEGLLHQLTPSQYQKIVDRNLILKEIYIGAMVIVLLVNSKLVY	177
QY	181	ANVTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFTLSQLGDAGKIKQVGIIICQEST	240
Db	178	ANVTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFTLSQLGDITTKIKQVGIIICQEST	237
QY	241	RIGDVKKVGYVDIDLLSAASKSPIIAEBEIHGAQPLDGVTFVLMSBGLYKALEAHP	300

Db	328	FRIDYKXKXNFNDI	ELLSTAKSKPI	TAPEEIHGCP	LDVGTGFLV	MSSEGLYKAL	ESAH	297								
Qy	301	GGQAND	ELIAMIDTE	FAKOTSL	DAVNO	AVDRYKR	IHSDFP	ASGGGRARCPHEMTL 360								
Db	298	GGQAND	ELIAMITE	FAKOVSL	DEVQAL	VERKRI	IHHDP	FASGGERAKYCSHEDWTL 357								
Qy	361	LVRFNG	VLGMSOPT	SPSPA	AGRV	PVSPV	PYSSA	QSTSKTSVTLSLWVPSOGOMWNG 420								
Db	358	LVRLNL	GPLG	IS	PFITL	--	PLQGRL	LYPVSPV	PSSAQN	SKISVTL	SLWVPS	GGPWNG 415				
Qy	421	AHAS	STLDEAT	PVLL	TNOSP	PTLT	TLOST	NTHT	TOSSSS	SSDGL	FRSR	PAHSL	PGEDGR	VEP 480		
Db	416	TN	SSSTL	DGTSTL	--	OSP	SATL	OSTW	HTHT	TOSSSS	SSDGL	FRSR	PLP	SLQPD	EDGR	VEP 473
Qy	481	YVDF	FAE	FR	LM	SV	DHGE	--	SVYTA 503							
Db	474	YVDF	TDF	FR	LM	MAEH	NDP	GT	LLTA 497							

Db	328	FRIDYKXKXNFNDI	ELLSTAKSKPI	TAPEEIHGCP	LDVGTGFLV	MSSEGLYKAL	ESAH	297								
Qy	301	GGQAND	ELIAMIDTE	FAKOTSL	DAVNO	AVDRYKR	IHSDFP	ASGGGRARCPHEMTL 360								
Db	298	GGQAND	ELIAMITE	FAKOVSL	DEVQAL	VERKRI	IHHDP	FASGGERAKYCSHEDWTL 357								
Qy	361	LVRFNG	VLGMSOPT	SPSPA	AGRV	PVSPV	PYSSA	QSTSKTSVTLSLWVPSOGOMWNG 420								
Db	358	LVRLNL	GPLG	IS	PFITL	--	PLQGRL	LYPVSPV	PSSAQN	SKISVTL	SLWVPS	GGPWNG 415				
Qy	421	AHAS	STLDEAT	PVLL	TNOSP	PTLT	TLOST	NTHT	TOSSSS	SSDGL	FRSR	PAHSL	PGEDGR	VEP 480		
Db	416	TN	SSSTL	DGTSTL	--	OSP	SATL	OSTW	HTHT	TOSSSS	SSDGL	FRSR	PLP	SLQPD	EDGR	VEP 473
Qy	481	YVDF	FAE	FR	LM	SV	DHGE	--	SVYTA 503							
Db	474	YVDF	TDF	FR	LM	MAEH	NDP	GT	LLTA 497							

Db	328	FRIDYKXKXNFNDI	ELLSTAKSKPI	TAPEEIHGCP	LDVGTGFLV	MSSEGLYKAL	ESAH	297								
Qy	301	GGQAND	ELIAMIDTE	FAKOTSL	DAVNO	AVDRYKR	IHSDFP	ASGGGRARCPHEMTL 360								
Db	298	GGQAND	ELIAMITE	FAKOVSL	DEVQAL	VERKRI	IHHDP	FASGGERAKYCSHEDWTL 357								
Qy	361	LVRFNG	VLGMSOPT	SPSPA	AGRV	PVSPV	PYSSA	QSTSKTSVTLSLWVPSOGOMWNG 420								
Db	358	LVRLNL	GPLG	IS	PFITL	--	PLQGRL	LYPVSPV	PSSAQN	SKISVTL	SLWVPS	GGPWNG 415				
Qy	421	AHAS	STLDEAT	PVLL	TNOSP	PTLT	TLOST	NTHT	TOSSSS	SSDGL	FRSR	PAHSL	PGEDGR	VEP 480		
Db	416	TN	SSSTL	DGTSTL	--	OSP	SATL	OSTW	HTHT	TOSSSS	SSDGL	FRSR	PLP	SLQPD	EDGR	VEP 473
Qy	481	YVDF	FAE	FR	LM	SV	DHGE	--	SVYTA 503							
Db	474	YVDF	TDF	FR	LM	MAEH	NDP	GT	LLTA 497							

Db	328	FRIDYKXKXNFNDI	ELLSTAKSKPI	TAPEEIHGCP	LDVGTGFLV	MSSEGLYKAL	ESAH	297								
Qy	301	GGQAND	ELIAMIDTE	FAKOTSL	DAVNO	AVDRYKR	IHSDFP	ASGGGRARCPHEMTL 360								
Db	298	GGQAND	ELIAMITE	FAKOVSL	DEVQAL	VERKRI	IHHDP	FASGGERAKYCSHEDWTL 357								
Qy	361	LVRFNG	VLGMSOPT	SPSPA	AGRV	PVSPV	PYSSA	QSTSKTSVTLSLWVPSOGOMWNG 420								
Db	358	LVRLNL	GPLG	IS	PFITL	--	PLQGRL	LYPVSPV	PSSAQN	SKISVTL	SLWVPS	GGPWNG 415				
Qy	421	AHAS	STLDEAT	PVLL	TNOSP	PTLT	TLOST	NTHT	TOSSSS	SSDGL	FRSR	PAHSL	PGEDGR	VEP 480		
Db	416	TN	SSSTL	DGTSTL	--	OSP	SATL	OSTW	HTHT	TOSSSS	SSDGL	FRSR	PLP	SLQPD	EDGR	VEP 473
Qy	481	YVDF	FAE	FR	LM	SV	DHGE	--	SVYTA 503							
Db	474	YVDF	TDF	FR	LM	MAEH	NDP	GT	LLTA 497							

Db	328	FRIDYKXKXNFNDI	ELLSTAKSKPI	TAPEEIHGCP	LDVGTGFLV	MSSEGLYKAL	ESAH	297				
Qy	301	GGQAND	ELIAMIDTE	FAKOTSL	DAVNO	AVDRYKR	IHSDFP	ASGGGRARCPHEMTL 360				
Db	298	GGQAND	ELIAMITE	FAKOVSL	DEVQAL	VERKRI	IHHDP	FASGGERAKYCSHEDWTL 357				
Qy	361	LVRFNG	VLGMSOPT	SPSPA	AGRV	PVSPV	PYSSA	QSTSKTSVTLSLWVPSOGOMWNG 420				
Db	358	LVRLNL	GPLG	IS	PFITL	--	PLQGRL	LYPVSPV	PSSAQN	SKISVTL	SLWVPS	GGPWNG 415


```

DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Protein phosphatase 2C.
MP09.
OS Mesembryanthemum crystallinum (Common ice plant).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
OX NCBI_TaxID=3544;
RN [1]
RP SEQUENCE FROM N.A.
RA Miyazaki S., Izumi S., Fukuhara T., Bohner H.J.;
RT "Plant Protein Phosphatases 2C? A large protein family serving
RT complex functions.";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083482; BAB88944.1; -
SQ SEQUENCE 319 AA; 34477 MW; 0BD228603E34CFB0 CRC64;

Query Match 7.9%; Score 205; DB 10; Length 319;
Best Local Similarity 23.1%; Pred. No. 1.9e-06;
Matches 83; Conservative 61; Mismatches 109; Indels 106; Gaps 14;

QY 12 EQQPSWTDDLPCHLSGVGSASNRYSADGKTESHPEDSMKFRSNN-----CF 63
DB 17 EEQP-----VSGGIGISQNGKFS--YGVASSPGKRSMSMEDFYETRIDGVEGEVVG 63
QY 64 LYGVENGYDGNVTNFWQORLSAEILLGOLNAEHA--ADYRVLLQAFDVERSFILESI 121
DB 64 LFGVDDGHGGAARAEVYKQNLFSNLI-----KHPKFSIDTKSAIAEYVTHDSEFLKSE 117
QY 122 DDALAERKASLQSLPEGVPOHQLPPQYOKIERLKLTEREISGMAVAVVNLNNKLYVA 181
DB 118 N-----TQNRD--AGSTASTALLVGRLLVA 141
QY 182 NGNTRALLCKSTVQGLQVLTQNLVDHTTENEDELFRLSQLG---LDACKIKQVIGIQGE 238
DB 142 NVGDSRAVTCR---GGEIAVSRDHKPKQSDERQRIEDAGSFVWAGTWKVGVLAA--- 194
QY 239 STRRIGDYKVKYGYTDIDLLSAKSKPPIAEPEIHGAQPLDVGTVGLVMEGLYKALEA 298
DB 195 VSRAPGDKLTK-----QYVADBEIQ--EEVVDSSLEFLLASDGLMDVVTN 239
QY 299 AHG-----PGQANQETAMITFEAKQTSLDVAQAIV-----DRYKRHSHT 341
DB 240 EEAVTWVKPIQDTBEAKAKLMQEAAYQSGSADNITCVVAFRLDNPIDSSNRSIGNHGT 298

RESULT 7
Q9LUS8 PRELIMINARY; PRT; 493 AA.
ID Q9LUS8;
AC Q9LUS8;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Similarity to protein phosphatase-2c.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Tabata S.;
RT Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RX MEDLINE=20277480; PubMed=10819329;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence
RT features of the regions of 4,504,864 bp covered by sixty P1 and TAC

```

```

RT clones";
RL DNA Res. 7:131-135(2000).
DR EMBL; AB022217; BAB02747.1.-
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2C; 1.
DR SMART; SM00331; PP2C SIG; 1.
SQ SEQUENCE 493 AA; 53613 MW; 613ED80B06B8C844 CRC64;

Query Match 7.8%; Score 202; DB 10; Length 493;
Best Local Similarity 23.5%; Pred. No. 6.2e-06;
Matches 96; Conservative 59; Mismatches 133; Indels 120; Gaps 18;

QY 24 CHLSGVGSAS--NRYSADGKTE---SHPPEDSMKFRSENNCFLY--GVENGYDGNVT 77
DB 129 CYLTSSSPSSSVNEGFLSAMEVQVAGAGDVRQAVCSFENGMVFCALYDGFNGRDAA 188
QY 78 NFVAQRLL-----SAELLGOLNAEHAED-----VR 103
DB 189 DFLACTLYESIVPHQLDRLQKQKTSDDGKELLSINISVDYSSTDLFRQGLVDCIN 248
QY 104 RVLLQAFDVERSFILESIDDALAEKASLQSLPEGVPOHQLPPQYOKIERLKLTEREIS 163
DB 249 RALFQA-----ETDFLRVMEQEMEERPD-----VS 274
QY 164 GGMVAVVAVVNLNNKLYAVAVGTRNALLCKSTVDG---LQVLTQNLVDHTTENEDELFRLSQ 220
DB 275 VGSQVAVTLLVGRDLYVNLIGDSRAVL--ATYNGNKQLQAVQLTEDHTVDNEVEARLLS 332
QY 221 LGSADGKIKQVIGIQGQ--ESTRIRIGDYKVKYGYTDIDLLSA-----KSPPIA- 268
DB 333 EHLDDPKIYVIGKIKKLKLVTRALG-----VGLKKKELNDALMGLRYRNLSPPYVS 387
QY 269 EP--EIHGAQPLDVGTVGLVLMSEGLYKALEAHGEGQANQETAMITFEAKQTSLD-- 324
DB 388 EPGMRVAKITIEDS--HFYIVASDGLFDFP-----SNEBAIGLVHSFVSNPSGCPA 436
QY 325 -----AVQAQVVDYKRIHSDTFPAGGERARFCPRHEMTLLVNF 366
DB 437 KILLERLVAQAARAGFTLBEITVPAGRKKRY--HDDVITVITL 481

RESULT 8
Q9S9Z7 PRELIMINARY; PRT; 282 AA.
ID Q9S9Z7;
AC Q9S9Z7;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE P21H2.4 protein (protein phosphatase type 2C-like protein).
GN P21H2.4.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC Vysotskaya V.S., Schwartz J.R., Yu G., Toriumi M., Lenz C., Liu S.,
RA Lee J., Liu A., Li J., Kremenetskaya I., Luros J., Gonzalez A.,
RA Altafi H., Araujo R., Brooks S., Buehler E., Chao O., Conn L.,
RA Conway A.B., Dunn P., Hansen N., Huizer L., Khan S., Kim C., Palm C.,
RA Rowley D., Shim P., Walker M., Davis R.W., Ecker J.R.;
RA Federspiel N.A.; Theologis A.;
RT "The sequence of BAC P21H2.4 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA STRAIN=CV. COLUMBIA;
RC Theologis A.;
RA Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

```

```
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Chen H., Cheuk R.,
RA Chung M.K., Ecker J., Theologis A., Davis R.W.;
RA Yamada K., Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007894; AAD46006.1; -
DR EMBL; AF370608; AAK43927.1; -
DR HSSP; P35813; 1A6Q.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
SQ SEQUENCE 282 AA; 30984 MW; 36404709FB03DA8D CRC64;

Query Match      7.5%; Score 194.5; DB 10; Length 282;
Best Local Similarity 25.4%; Pred. No. 8.7e-06;
Matches 86; Conservative 45; Mismatches 97; Indels 111; Gaps 15;

QY 29 VCSASNRS-----YSAG-----KGTESHPPED-----SWLKFRSENNCFYGVNFGYD 72
Db 16 VGRASTSGKGRNNDGEIKFGVSLYKGRANHPMEDYHVSKEFKI-DGNELGLFAIYDGH 74
QY 73 GNRVTFNFAQRLSAELLQLQNAEADVRVLLQAFDVVERSFLSIDDLAEKASLQ 132
Db 75 GERVAYLOKHLFSNI-----LKEGFYRDPQKSIIAAEKTD----- 112
QY 133 SOLPEGVQHQLPPOYQKILERLKLTEREISGGANAVAVLLN-NKLYVANVGNRALLC 191
Db 113 -----QAILSHSSDLGR---GGSTAVTATLNGRRLLVWVNGDSRAVL 153
QY 192 KSTVDGLQVTLQNVHDHTTENEDELFRLSQLGLDA-----GKIQVGIICQESTRIGD 245
Db 154 Q-----GGQAQMTIDHEPHT-----RLSTEGKGFVSNMGPVPRVN-----GQ 194
QY 246 YKVKYGYTDIDLSSAKSPKPIAEPIHGAQPLDGVTCFLVLMSEGLYKALEAAHFGQQA 305
Db 195 LAVSRAFGDKSLKTHRSDPDVKQSSI-----DDHTDVLVLSADGLWKVM-----A 240
QY 306 NOE-----IAAMIDTEPAKQTSLDVAQAQVV 331
Db 241 NQEAIDARIIRIKPLKAAKELTTEALRDSKDDISCIV 279

RESULT 9
Q9LNF4
ID Q9LNF4 PRELIMINARY; PRT; 377 AA.
AC Q9LNF4
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE F21D18.27 (Protein phosphatase-2C, putative).
GN T2J15.5.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
RA Shinn P., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
RA Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N., Nguyen M.,
RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
RA Toriumi M., Vaysberg M., Yu G., Federspiel N.A., Theologis A.,
RA Ecker J.R.;
RT "Genomic sequence for Arabidopsis thaliana BAC F21D18 from chromosome
RT I.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RP Ecker J.R.;
```

```
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA Ecker J.R.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE FROM N.A.
RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E.,
RA Conn L., Conway A., Gonzales A., Hansen N., Howing B., Koo T., Lam B.,
RA Lee J., Lenz C., Li J., Liu A., Liu K., Liu S., Mukharsky N.,
RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
RA Theologis A., Ecker J.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Egu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzer S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:816-820(2000).
DR EMBL; AC023673; AAF79528.1; -
DR EMBL; AC051631; AAG51521.1; -
DR HSSP; P35813; 1A6Q.
DR InterPro; IPR000222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS01032; PP2C; 1.
SQ SEQUENCE 377 AA; 41431 MW; 586D46246FB64F5F CRC64;

Query Match      7.5%; Score 193.5; DB 10; Length 377;
Best Local Similarity 24.9%; Pred. No. 1.6e-05;
Matches 77; Conservative 47; Mismatches 114; Indels 71; Gaps 11;

QY 65 YGVNFGYDGNRVNFMVAQRLSAELLQLQNAEADVRVLLQAFDVVERSFLSIDD- 123
Db 111 YGVFDGHHGPEAIAFMKENLT-----RLFFQDA-----VFPMPISIVDAFFLELENS 158
QY 124 -----ALAEKASLSQLPPEGVQHQLPPOYQKILERLKLTEREISGGANAVAVLLNKL 178
Db 159 HRKAFALADLAWADETIVSG-----SCGTTALTALIIGHRL 194
QY 179 YVANVGNRALLCKSTVDGLQVTLQNVHDHTTENEDELFRLSQLG--LDAGKIKQVGIICG 236
Db 195 LVANAGDCRAVLCRCGV-----AVDMSFDRHSYFERRRIEDLGGYFEDGYUNGVLAV-- 248
QY 237 QESTRIGDYKVKYGYTDIDLSSAKSPKPIAEPIHGAQPLDGVTCFLVLMSEGLYKAL 296
Db 249 ---TRAIGDWELKNPPTD-----SSSLISDPBI-GQILTDEDFLILACDGIWDVL 297
QY 297 EAAHFGQQAQ-----ETAMIDTEFAKQTSLDVAQAQVDRVRIHSDTFASGGE 347
Db 298 SSQNAVSNVRQGLRRHGDPRQAMELGKAEARLQSSDNMT-VIVICFSSVPSPKQPRR 356
```

Y	348	RARPCPRHE	356
Db	357	RURCVSDE	365
RESULT	10		
ID	081716	PRELIMINARY;	PRT; 355 AA.
AC	081716;		
DT	01-NOV-1998	(TrEMBLrel. 08, Created)	
DT	01-NOV-1998	(TrEMBLrel. 08, Last sequence update)	
DT	01-JUN-2002	(TrEMBLrel. 21, Last annotation update)	
DE	Hydrotheoretical 39.4 kDa protein (Protein phosphatase 2C).		
GN	ATPG25070 OR ATPPC4.2		
OS	Arabidopsis thaliana (mouse-ear cress).		
OC	Euparietales; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;		
OC	euprosids II; Brassicales; Brassicaceae; Arabidopsis.		
CC	NCBI_TaxId=3702;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, COLUMBIA;		
RX	MEDLINE=20088487; PubMed=10617197;		
RA	Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,		
RA	Fuji C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,		
RA	Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.		
RA	Cionin L.A., Shen M., Vanden S.E., Umayam L., Tallon L.J., Gill J.E.		
RA	Adams M.D., Carreira A.J., Creasy T.H., Goodman H.M., Somerville C.R.,		
RA	Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,		
RA	Salzberg S.L., Fraser C.M., Venter J.C.;		
RT	"Sequence and analysis of chromosome 2 of the plant Arabidopsis		
RT	thaliana.";		
RL	Nature 402:761-768 (1999).		
RP	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=CV, COLUMBIA;		
RA	Lin X.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RP	[3]		
RP	SEQUENCE FROM N.A.		
RA	Yamada K., Liu S.X., Pham P.K., Banh J., Dale J.M., Goldsmith A.D.,		
RA	Jiang P.X., Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M.,		
RA	Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,		
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,		
RA	Kawai J., Kim C., Koeseena E., Lam B., Lin J., Meyers M.C., Miranda M.,		
RA	Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,		
RA	Shin P., Southwick A., Tracy S.E., Shinozaki K., Davis R.W.,		
RA	Ecker J.R., Theologis A.;		
RT	"Full length cDNA of gene F27G12.1/Atg25070 (GI_4559345).";		
RT	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
RP	[4]		
RP	SEQUENCE FROM N.A.		
RA	Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,		
RA	Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,		
RA	Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,		
RA	Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,		
RA	Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,		
RA	Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,		
RA	Sakurai T., Satou M., Seki M., Shin P., Southwick A., Shinozaki K.,		
RA	Davis R.W., Ecker J.R., Theologis A.;		
RT	"Arabidopsis Open Reading Frame (ORF) Clones.";		
RT	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.		
RP	[5]		
RP	SEQUENCE FROM N.A.		
RA	Izumi S., Yamada M., Onato H., Miyazaki S., Bohnert H.J.,		
RA	Fukuhara T.;		
RT	"Substrate Specificity of Type 2C Protein Phosphatases (PP2C) in		
RT	Arabidopsis thaliana.";		
RT	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AC006565; AAD23006.1; -		
DR	EMBL; AY050873; AAK92810.1; -		
DR	EMBL; AY091209; AAM4148.1; -		
DR	EMBL; AB079671; BAB84700.1; -		

```

DR HSSP: P35813; 1A60.
DR InterPro; IPR000222; PP2C.
DR InterPro; IPR001934; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C SIG; 1.
DR PROSITE; PS01033; PP2C; 1.
DR Hypothetical protein.
KW SEQUENCE 355 AA; 39354 MW; CAD638796203C746 CRC64;
SQ
Query Match 7.5%; Score 192.5; DB 10; Length 355;
Best Local Similarity 24.0%; Pred. No. 1.7e-05;
Matches 82; Conservative 61; Mismatches 128; Indels 71; Gaps 16;
OY YGVENGIDGNVYTFVVAQRSLSAELLGLQ-LNAEAEADVRVLLQADVV--ERSFLEST 121
D 53 FGVIYGHGKGVAAVFCCKAYTLHQVIVSNEAYKTDGVTSLRAKFRMDMNOGGMRELA 112
OY 122 DDALAEKASLSQSLPEGV-----PQHLRPOVQKILERLKTIERISG- 164
D 113 --VLGGKNNKSSGMIIEGFIWSPRGSDTNNQDPSWPLEDGP-----HSDTGP 157
OY 165 --GAAVAVAVLNNKLYYANVGTNRALLCKSTVGLQVTLQNLVDTTENEDELRLSQLG 222
D 158 TSGCACAVALLKDKKLFLPANNAGDSRCVISRKS---QAVNLSKHKRDLEVEKERI---- 209
OY 223 LDAGKIKOVGIICQO-ESTRIRIGYKVKYGYTIDILISAASKRITLAEPIHGAQPLDGV 281
D 210 LKAGGFIFHAGRIINGSLNLTIRAIIGMEFQKQ---NKLFPSEKQVWTAADPINTIDLCD- 263
OY 282 TGFVLIMSEGIYKALEAHGPGCANOEFLAAMIDIEFPAKQTSGLDAVAVDVRYARIRISDT 341
D 264 DDFLVAACDGIWDCI-----SSQELVDFTIHQULKSETKSLSTVCEKVDKRC--LAPDT 313
OY 342 FASGGERARFCPRHEDMTLLVRNFGYPLGEMSQEPTPS-PAPA 382
D 314 ATGEG-----C---DNMTIILVQFKKPPSPSTBEDSDKPEES 347
RESULT 11
OBS3P1 PRELIMINARY; PRT; 243 AA.
ID AC OBS3P1
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 26.7 kDa protein.
GN 24K23.16.
OS Oryza sativa (Japanese cultivated group).
OC Eukarya; eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRKIN=CV. NIPONBARE;
RA Park Y.-U., Kostecka N., Ramakrishna W., Samiuel P., Shiloff B.,
RA Ma J., Jiang Z., Kleinofs A., Bennett J.;
RT "Sequence characterization of orthologous regions in the barley and
RT rice genomes.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF480497; AAL67187.1; -.
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 26734 MW; 8B481E17F559CE5F CRC64;
Query Match 7.4%; Score 191.5; DB 10; Length 243;
Best Local Similarity 24.8%; Pred. No. 1.1e-05;
Matches 70; Conservative 47; Mismatches 82; Indels 83; Gaps 11;
OY 64 LVGVNGYDGNKRVTFVVAQRSLSAELLGLQALAEHAE--ADVRRVLLQAFDVERSFLEST 121
D 19 LFGVYDGHGKGVAAVVAQNLFSHLL-----RHPKFIQDTKVAIDAYKSTDSDFLES- 71

```



```

Db 111 -----OKLIANDRTDLE---SGGSTAVTALLINGKALMIANVGDSPRAIVSRS 155
QY 195 VDLGQVLTQNVDTTENEDELFRLSQGLDAGKIKQVGIIGQESTRIIDGYKVKYGYTD 254
Db 156 ----KAKQMSVDHDDDDPTERSMIESKG-----GFV-----TNRPGVPRVNGILLA 197
QY 235 IDLLSAASKP--IIAPEIHQAQPLDGVGTGLVMSGLVKALEAANGPOANOELIAM 312
Db 198 ISSVGDKKLKAYLNSEPEIKDV-TIDSHTDPIIVASDGIKVM-----SNOE---A 245
QY 313 IDTEFAKQTSLDVAQAQVAVDRVKRIHSD 340
Db 246 DVAKKLDKPKELARQVVAEALKNRNSKD 273

RESULT 14
Q9LEW5 PRELIMINARY; PRT; 348 AA.
AC Q9LEW5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Protein phosphatase 2C-like protein.
GN T30N20.10.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL365234; CAB96829.1; -.
DR HSSP; P35813; IAGQ.
DR InterPro; IPR001993; Mitoch_carrier.
DR InterPro; IPR000222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE; PS01032; PP2C; 1.
SQ SEQUENCE 348 AA; 37379 MW; 12447BB07F740E2C CRC64;

Query Match 7.4%; Score 190.5; DB 10; Length 348;
Best Local Similarity 20.2%; Pred. No. 2,3e-05;
Matches 94; Conservative 75; Mismatches 155; Indels 141; Gaps 18;

```

```

QY 302 PQANOELIAMT-----DTEFAKQTSLDVAQAQVAVDRVKRIHSDTEFAKQTSRERARFCPRHED 357
Db 232 ---SNEAAVAMVEKEDPE-----DSAKKLVEGKIRKGSAD-----N 265
QY 358 MTLVNFPGYPLGEMSQPTSPAPAGRVYPSVPSAOSTSKTSVTLSLVMSQGM 417
Db 266 ITCVVVRF-----LEKSSASSHISSSSSSEAKEMPEPLGDL 301
QY 418 VNGASHASTID---EATPLINQSTPLTLQSTNTHQSSSSSD 458
Db 302 AISSNEAKQVQIGSGNKPENVTNRKPDITASRSTDTLTERNSVTD 346

RESULT 15
Q9SZN2 PRELIMINARY; PRT; 395 AA.
AC Q9SZN2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 43.6 kDa protein.
GN P20M13.80 OR A74G38520.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxId=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Medler H., Kutzner M., Wambutt R., Bancroft I., Mewes H.W.,
RA Mayer K.F.X., Schueller C.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Medler H., Kutzner M., Wambutt R., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035540; CAB37508.1; -.
DR EMBL; AL161593; CAB80516.1; -.
DR InterPro; IPR000222; PP2C.
DR InterPro; IPR001932; PP2C-like.
DR Pfam; PF00481; PP2C; 1.
DR SMART; SM00332; PP2Cc; 1.
DR SMART; SM00331; PP2C_SIG; 1.
DR PROSITE; PS01032; PP2C; 1.
KM Hypothetical protein.
SQ SEQUENCE 395 AA; 43621 MW; 438B4906AFA4C2A CRC64;

Query Match 7.4%; Score 190.5; DB 10; Length 395;
Best Local Similarity 24.5%; Pred. No. 2,8e-05;
Matches 104; Conservative 60; Mismatches 150; Indels 111; Gaps 20;

```


Search completed: December 9, 2002, 22:58:26
Job time : 58.4488 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 9, 2002, 22:48:29 ; Search time 21.4072 Seconds
(without alignments)
692.718 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRRLQLSQPQSWTDD.....AEFYRLMSVDHGQSWTAP 504

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/1aa/5A-COMB.pep: *
2: /cgn2_6/prodata/1/1aa/5B-COMB.pep: *
3: /cgn2_6/prodata/1/1aa/6A-COMB.pep: *
4: /cgn2_6/prodata/1/1aa/6B-COMB.pep: *
5: /cgn2_6/prodata/1/1aa/PCITUS-COMB.pep: *
6: /cgn2_6/prodata/1/1aa/backfiles1.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2580	100.0	504	2	US-08-752-891-2
2	2580	100.0	504	2	US-09-144-178-2
3	2580	100.0	504	2	US-09-406-854-2
4	2580	100.0	504	4	US-09-529-279-2
5	2580	100.0	517	4	US-09-529-279-11
6	2577	99.9	513	4	US-09-529-279-43
7	2575	99.8	504	2	US-08-752-891-6
8	2575	99.8	504	2	US-09-144-178-6
9	2575	99.8	504	4	US-09-406-854-6
10	178	6.9	392	4	US-09-013-881-2
11	125.5	4.9	346	3	US-08-935-855-20
12	124	4.8	306	2	US-08-822-701-8
13	124	4.8	306	3	US-08-935-855-8
14	118	4.6	390	2	US-09-206-646-4
15	118	4.6	392	2	US-08-822-701-2
16	118	4.6	392	3	US-08-935-855-2
17	118	4.6	542	3	US-08-935-855-22
18	117.5	4.6	309	2	US-08-822-701-7
19	117.5	4.6	309	3	US-08-935-855-7
20	117	4.5	390	2	US-08-873-093-3
21	117	4.5	390	4	US-09-206-646-3
22	117	4.5	677	1	US-08-188-582-13
23	117	4.5	677	1	US-08-646-715-13
24	116	4.5	314	2	US-08-822-701-10
25	116	4.5	314	3	US-08-935-855-10
26	110	4.3	479	4	US-09-206-646-1
27	109	4.2	478	2	US-08-873-093-1

28	109	4.2	478	2	US-08-873-093-4	Sequence 4, Appli
29	108	4.2	281	2	US-08-822-701-9	Sequence 9, Appli
30	108	4.2	281	3	US-08-935-855-9	Sequence 9, Appli
31	106	4.1	910	4	US-08-460-269C-2	Sequence 2, Appli
32	105.5	4.1	2930	4	US-09-417-822-2	Sequence 2, Appli
33	105	4.1	1068	4	US-09-085-199B-11	Sequence 11, Appl
34	104.5	4.1	3724	2	US-08-804-227C-10	Sequence 10, Appl
35	104.5	4.1	3724	2	US-08-804-198-4	Sequence 4, Appli
36	103.5	4.0	1066	4	US-09-541-782-8	Sequence 8, Appli
37	103.5	4.0	1066	4	US-09-723-820-8	Sequence 8, Appli
38	103.5	4.0	2471	1	US-08-185-432-16	Sequence 16, Appl
39	103.5	4.0	2471	1	US-08-083-590A-19	Sequence 19, Appl
40	103.5	4.0	2471	3	US-08-532-384-19	Sequence 19, Appl
41	103.5	4.0	2471	4	US-08-899-232-1	Sequence 1, Appli
42	102.5	4.0	594	2	US-08-910-856-2	Sequence 2, Appli
43	100.5	3.9	441	4	US-09-522-666-4	Sequence 4, Appli
44	100.5	3.9	1024	4	US-09-522-666-2	Sequence 2, Appli
45	100.5	3.9	1049	4	US-09-522-666-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-752-891-2
; Sequence 2, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-752-891-2

Query Match 100.0%; Score 2580; DB 2; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MAAORSLTSEOQPSWMTDDLPLCHLSGVGSASNRXSVAQCKGTESHPPEDSMJLKRSEN	60
Db	1	MAAORSLTSEOQPSWMTDDLPLCHLSGVGSASNRXSVAQCKGTESHPPEDSMJLKRSEN	60
QY	61	NCFLYGVENGVDGNRVTFNFAQRLSABELLIGOLMAEHAEDVRRVLLQAFDVERSFLES	120
Db	61	NCFLYGVENGVDGNRVTFNFAQRLSABELLIGOLMAEHAEDVRRVLLQAFDVERSFLES	120
QY	121	IDDDLAERASIQOSLPBGPPOHOLPPOYOKILERLKTLEREISGGAAVAVAVLNNKLXY	180
Db	121	IDDDLAERASIQOSLPBGPPOHOLPPOYOKILERLKTLEREISGGAAVAVAVLNNKLXY	180
QY	181	ANVCNRRALLCKSTVDGLQVTLQNLVNDHTTENEDFLRSLQGLDAGKIKOVGIIICQOEST	240
Db	181	ANVCNRRALLCKSTVDGLQVTLQNLVNDHTTENEDFLRSLQGLDAGKIKOVGIIICQOEST	240
QY	241	RRIQGYKKYCYTDDILLSAKSKPIIAEPRIHAQPLDGVGTGLVYMSEGLKALEAAH	300
Db	241	RRIQGYKKYCYTDDILLSAKSKPIIAEPRIHAQPLDGVGTGLVYMSEGLKALEAAH	300
QY	301	GGQANDOEIAAMIDTEFAKQTSLDVAQAQVVDYKRIHSDTFASGGERAPCPRHEDMTL	360
Db	301	GGQANDOEIAAMIDTEFAKQTSLDVAQAQVVDYKRIHSDTFASGGERAPCPRHEDMTL	360
QY	361	LVRNNGYTLGEMSOPTBPAPAAAGRYVYVPSASQSTSKTSVTLSLVMPSGQCMVNG	420
Db	361	LVRNNGYTLGEMSOPTBPAPAAAGRYVYVPSASQSTSKTSVTLSLVMPSGQCMVNG	420
QY	421	AHSASTLDEAFPTLLNQSPTLLQSTTNHTQSSSSSSDGLFRSRPAHSLPPGEGDVERP	480
Db	421	AHSASTLDEAFPTLLNQSPTLLQSTTNHTQSSSSSSDGLFRSRPAHSLPPGEGDVERP	480
QY	481	YVDFAEFRLMSVHGEQSVVTAP	504
Db	481	YVDFAEFRLMSVHGEQSVVTAP	504

RESULT 2
US-09-144-178-2
Sequence 2, Application US/09144178
Patent No. 5989862
GENERAL INFORMATION:
APPLICANT: MATSUMOTO, Kunihito
APPLICANT: NISHIDA, Eisuke
TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,178
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.

```

?      REGISTRATION NUMBER: 29,768
?      REFERENCE/DOCKET NUMBER: 17988/111
?      TELECOMMUNICATION INFORMATION:
?      TELEPHONE: (202)672-5300
?      TELEFAX: (202)672-5399
?      TELEX: 904136
?      INFORMATION FOR SEQ ID NO: 2 :
?      SEQUENCE CHARACTERISTICS:
?          LENGTH: 504 amino acids
?              type: amino acid
?              TOPOLOGY: linear
?      MOLECULE TYPE: protein
?
US-09-144-178-2

```

Query Match	100.0%;	Score 2580;	DB 2;	Length 504;
Best Local Similarity	100.0%;	Pred. No.9,6e-230;		
Matches 504; Conservative	0;	Mismatches	0;	Gaps 0

QY	1	MAORRSLLOEOPQSMYTDPLPLCHLSGVSSASNRSTYSADGKOTESHPPEDSWLKJRSEN	60
Db	1	MAORRSLLOEQQPSMTDDPLPLCHLSGVSSASNRSTYSADGKOTESHPPEDSWLKJRSEN	60
QY	61	NCFLYGVFNQYDGNRVTNFVAQRLSABLLIGOLNAEHAEDVRVLLQAFDVERSFLES	120
Db	61	NCFLYGVFNQYDGNRVTNFVAQRLSABLLIGOLNAEHAEDVRVLLQAFDVERSFLES	120
QY	121	IDDLALAEKASLQSQLPBGVPOHOLPPOYOKILLEBLKTLEREISGGAMAVALVLLNNKLY	180
Db	121	IDDLALAEKASLQSQLPBGVPOHOLPPOYOKILLEBLKTLEREISGGAMAVALVLLNNKLY	180
QY	181	ANVTNRALLCKSTVDGQVTVLQNLVNDHTTENEDFLRLSQLGLDAGKIKOVGII	240
Db	181	ANVTNRALLCKSTVDGQVTVLQNLVNDHTTENEDFLRLSQLGLDAGKIKOVGII	240
QY	241	RIRIGYKVKYCYTDILDLSSAKSKPIIAEPHGAQPLDQVTFGLVYMSGLKKALEAAH	300
Db	241	RIRIGYKVKYCTDILDLSSAKSKPIIAEPHGAQPLDQVTFGLVYMSGLKKALEAAH	300
QY	301	GGQANQEIAMIDTEFAKQTSLDVAQAQAVDVRKYHSDTFASSGGERAFPCRHEMDTL	360
Db	301	GGQANQEIAMIPTTEFAKQTSLDVAQAQAVDVRKYHSDTFASSGGERAFPCRHEMDTL	360
QY	361	LVNRNGYGLGMSOPTSPAPAAAGRYYPVSPVSSAQOSTSKISVTLLSLMPGOGQWNG	420
Db	361	LVNRNGYGLGMSOPTSPAPAAAGRYYPVSPVSSAQOSTSKISVTLLSLMPGOGQWNG	420
QY	421	AHSASTLDEATPTLLNOSPVTLLIOSTVHTTOSSSSSSDGGRFRGRPHSLPPGEDGRVEP	480
Db	421	AHSASTLDEATPTLLNOSPVTLLIOSTVHTTOSSSSSSDGGRFRGRPHSLPPGEDGRVEP	480
QY	481	YVDFAEFRYLRMSVDHGEOSVTPAR	504
Db	481	YVDFAEFRYLRMSVDHGEOSVTPAR	504

RESULT 3
 US-09-406-854-2
 ; Sequence 2, Application US/09406854
 ; Patent No. 6140042
 ;
 ; GENERAL INFORMATION:
 ;
 APPLICANT: MATSUMOTO, Kunihito
 APPLICANT: NISHIDA, Eisuke
 TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
 ;
 ; NUMBER OF SEQUENCES: 8
 ;
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Foley & Lardner
 STREET: 3000 K Street, N.W., Suite 500
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ;
 ; ZIP: 20007-5109
 ;
 ; COMPUTER READABLE FORM:
 ;
 ; MEDIUM TYPE: Floppy disk
 ;

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/406,854
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/752,891
FILING DATE: 20-NOV-1996
APPLICATION NUMBER: JP 8-300856
FILING DATE: 28-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 8-126282
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 17981/111
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 9041136
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 504 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-09-406-854-2

Query Match 100.0%; Score 2580; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSEN 60

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120

QY 121 IDDAEAKASLSQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
Db 121 IDDAEAKASLSQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180

QY 181 ANVGTNRALLCKSTVDGLQVTLNVHDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLNVHDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240

QY 241 RRGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300

QY 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360

QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420

QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 4

US-09-529-279-2
Sequence 2, Application US/09529279
Patent No. 6451617
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
FILE REFERENCE: 053466/0278
CURRENT APPLICATION NUMBER: US/09/529,279
CURRENT FILING DATE: 2000-04-11
PRIOR APPLICATION NUMBER: PCT/JP98/04796
PRIOR FILING DATE: 1998-10-22
PRIOR APPLICATION NUMBER: JP 9/290188
PRIOR FILING DATE: 1997-10-22
NUMBER OF SEQ ID NOS: 48
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 504
TYPE: PRT
ORGANISM: Homo sapiens
US-09-529-279-2

Query Match 100.0%; Score 2580; DB 4; Length 504;
Best Local Similarity 100.0%; Pred. No. 9.6e-230;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQPSWTDLPLCHLSGVGSASNRYSADGKTESHPEDSWLKFRSEN 60

QY 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVTFNFAQRLSAELLGQLNAEHAEDVRRVLLQAFDVVERSFLES 120

QY 121 IDDAEAKASLSQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180
Db 121 IDDAEAKASLSQSLPEGVPQHQLPPQYQKILRLKTLEREISGGAMAVVAVLLNNKLYV 180

QY 181 ANVGTNRALLCKSTVDGLQVTLNVHDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVTLNVHDHTTENEDELFRSLQGLDAGKIKQVGIICQEST 240

QY 241 RRGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRGDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300

QY 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQIEAAMIDTEFAKQTSLDAAVAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360

QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVSYSSAQSTSKTSVTLSLVMPSGQMWNG 420

QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHTQSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480

QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 5

US-09-529-279-11
Sequence 11, Application US/09529279
Patent No. 6451617
GENERAL INFORMATION:
APPLICANT: ONO, KOICHIRO
APPLICANT: OHTOMO, TOSHIHIKO
APPLICANT: TSUCHIYA, MASAYUKI
TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES

```
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-11

Query Match      100.0%; Score 2580; DB 4; Length 517;
Best Local Similarity 100.0%; Pred. No. 1e-229;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAQRSLLOSQEQPSWTDPLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
DB 1 MAQRSLLOSQEQPSWTDPLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
OY 61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 120
DB 61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 120
OY 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
DB 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
OY 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
DB 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
OY 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
DB 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
OY 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
DB 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
OY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
OY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
OY 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
OY 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
OY 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
DB 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
OY 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
DB 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
OY 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
DB 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
OY 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
DB 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
OY 481 YVDFAEFYRLMSVDHGQSVVTAP 504
DB 481 YVDFAEFYRLMSVDHGQSVVTAP 504

RESULT 6
US-09-529-279-43
; Sequence 43, Application US/09529279
; Patent No. 6451617
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/09/529,279
; CURRENT FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-529-279-43

Query Match      99.9%; Score 2577; DB 4; Length 513;
Best Local Similarity 99.8%; Pred. No. 1.9e-229;
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 MAQRSLLOSQEQPSWTDPLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
DB 1 MAQRSLLOSQEQPSWTDPLPLCHLSGVGSASNSYSADGKGTSHPPEDSWLKFRSEN 60
OY 61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 120
DB 61 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 120
OY 70 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 129
DB 70 NCFLYGVFNGYDGNRVTNFVAQRLSAELLGQLNAEADRVRLLOAFDVERSFILES 129
OY 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
DB 121 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 180
OY 130 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 189
DB 130 IDDALEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVLLNNKLYV 189
OY 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
DB 181 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 240
OY 190 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 249
DB 190 ANVTNRALLCKSTYVDGLQVTLNVDHTTENEDELFRLSQGLDAGKIKOVGIICQEST 249
OY 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
DB 241 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 300
OY 250 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 309
DB 250 RRIQDYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFVLWSEGLYKALEAAH 309
OY 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
OY 310 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 369
DB 310 GPGQANOETIAMIDTEFAKQTSIDAVAQAQVDRVKRIHSDTFASGGERARFCPRHEDMTL 369
OY 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
DB 361 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 420
OY 370 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 429
DB 370 LVNRFYPLGMSQPTSPAPAGGRVYVSVYSSAOSTSKTSYVTLSLVMSQGMVNG 429
OY 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
DB 421 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 480
OY 430 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 489
DB 430 AHSASTLDEATPTLTNOSPTLTLOSTNTHTOSSSSSDGGLFRSRPAHSLPGEEDGRVEP 489
OY 481 YVDFAEFYRLMSVDHGQSVVTAP 504
DB 481 YVDFAEFYRLMSVDHGQSVVTAP 504
OY 490 YVDFAEFYRLMSVDHGQSVVTAP 513
DB 490 YVDFAEFYRLMSVDHGQSVVTAP 513

RESULT 7
US-08-752-891-6
; Sequence 6, Application US/08752891
; Patent No. 5837819
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiko
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
```

```
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 8-300856
/ FILING DATE: 28-OCT-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 8-126282
/ FILING DATE: 24-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 17981/111
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)672-5300
/ TELEFAX: (202)672-5399
/ TELEX: 904136
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 504 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-752-891-6

Query Match          99.8%; Score 2575; DB 2; Length 504;
Best Local Similarity 99.8%; Pred. No. 2.8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLNAEHAEDVRRVLLQAPDVVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLNAEHAEDVRRVLLQAPDVVERSFLES 120
QY 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGAMAVVAVLLNNKLYV 180
Db 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
QY 241 RRTGDKYVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRTGDKYVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFRASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFRASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQGVWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQGVWNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480
Db 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480
QY 481 YVDFAEFYRLWSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLWSVDHGQSVVVTAP 504

RESULT 8
US-09-144-178-6
; Sequence 6, Application US/09144178
; Patent No. 5989862
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
```

```
/ ADDRESSEE: Foley & Lardner
/ STREET: 3000 K Street, N.W., Suite 500
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20007-5109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC Compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/144.178
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/752,891
/ FILING DATE: 20-NOV-1996
/ APPLICATION NUMBER: JP 8-300856
/ FILING DATE: 28-OCT-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: JP 8-126282
/ FILING DATE: 24-APR-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 17981/111
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202)672-5300
/ TELEFAX: (202)672-5399
/ TELEX: 904136
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 504 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-09-144-178-6

Query Match          99.8%; Score 2575; DB 2; Length 504;
Best Local Similarity 99.8%; Pred. No. 2.8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAAQRSLQSEQQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLNAEHAEDVRRVLLQAPDVVERSFLES 120
Db 61 NCFLYGVFNGYDGNRVNFTNFAQRLSAELLQLNAEHAEDVRRVLLQAPDVVERSFLES 120
QY 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGAMAVVAVLLNNKLYV 180
Db 121 IDDALEAKASLOSQLPBGVPOHQLPPQYQKILERLKTLEIREISGAMAVVAVLLNNKLYV 180
QY 181 ANVGTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
Db 181 ANVGTNRALLCKSTVDGLQVLTQNLNVDHTTENEDELFRLSQLGLDAGKIKQVGIIICQEST 240
QY 241 RRTGDKYVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
Db 241 RRTGDKYVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFRASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEIAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFRASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQGVWNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYVPVYSSAQSTSKTSVTLSLVMPSGQGVWNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480
Db 421 AHSASTLDEATPTLTNQSPTLTLOSTNTHQTQSSSSSDGGLFRSRPAHSLPPGEDGRVPE 480
```

Db 421 AHSASTLDEATPTLTNQSPTLTLLQSTNTHQTOSSSSDGGLFRSRPAHSLPPEGEDGRVER 480
QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

RESULT 9
US-09-406-854-6
; Sequence 6, Application US/09406854
; Patent No. 6140042
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihito
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 672-5300
; TELEFAX: (202) 672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-406-854-6

Query Match 99.8%; Score 2575; DB 4; Length 504;
Best Local Similarity 99.8%; Pred. No. 2,8e-229;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAQRSLLOSQEQPSWTDDLPICHSIGVSASNRYSADGKGTSHHPEDSWLKRSEN 60
Db 1 MAQRSLLOSQEQPSWTDDLPICHSIGVSASNRYSADGKGTSHHPEDSWLKRSEN 60
QY 61 NCFLYGVNFGYDGNRYTNFPAORLSAELLGOLNAEHAADYRVLQAFDVVERSFLES 120
Db 61 NCFLYGVNFGYDGNRYTNFPAORLSAELLGOLNAEHAADYRVLQAFDVVERSFLES 120
QY 121 IDBALAEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVALLNKKLV 180
Db 121 IDBALAEKASLOSQLEPGVPOHOLPPQYOKILERLKTLEREISGMAVAVALLNKKLV 180
QY 181 ANVTNRALLCKSTVDGLQVTLNVDDHTTENEDLFRLSQGLDAGKIKOVGIIIGQEST 240

Db 181 ANVTNRALLCKSTVDGLQVTLNVDDHTTENEDLFRLSQGLDAGKIKOVGIIIGQEST 240
QY 241 RRTGDYKVKYGVGTIDIDLSAAKSKPIAEPETHGQPLDGVTFVLVSEGLYKALEAAH 300
Db 241 RRTGDYKVKYGVGTIDIDLSAAKSKPIAEPETHGQPLDGVTFVLVSEGLYKALEAAH 300
QY 301 GPGQANQEIAMIDTEFAKQTSIDAVACA VVDVRKRIHSDTFASGGERARFCPRHEMTL 360
Db 301 GPGQANQEIAMIDTEFAKQTSIDAVACA VVDVRKRIHSDTFASGGERARFCPRHEMTL 360
QY 361 LVNFGYPIGEMSQPTSPAPAGGRVYVSPYSSAOSTKTSVTLSLVMSQGMWNG 420
Db 361 LVNFGYPIGEMSQPTSPAPAGGRVYVSPYSSAOSTKTSVTLSLVMSQGMWNG 420
QY 421 AHSASTLDEATPTLTNQSPTLTLLQSTNTHQTOSSSSDGGLFRSRPAHSLPPEGEDGRVER 480
Db 421 AHSASTLDEATPTLTNQSPTLTLLQSTNTHQTOSSSSDGGLFRSRPAHSLPPEGEDGRVER 480
QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
Db 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

RESULT 10
US-09-013-881-2
; Sequence 2, Application US/09013881
; Patent No. 6132964
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Lal, Preeti
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: HUMAN HYDROLASE-LIKE MOLECULES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,881
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: BILLINGS, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0470 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 392 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: KIDNOT02
; CLONE: 195647

US-09-013-891-2

Query Match 6.9%; Score 178; DB 4; Length 392;
Best Local Similarity 22.9%; Pred. No. 4.6e-08;
Matches 83; Conservative 65; Mismatches 116; Indels 98; Gaps 18;

QY 3 AQRRLQSEQQPSWTDL---FLCHLGVGSASNRYSADGKG-----T 44
DB 77 AKRKT---SEBEKNGSEELVEKKVKASSV-IFGLKGYVAERKGEREMQDAHVILNDIT 132
QY 45 ESHPEDSWLKRSENENFLYGVNGYDGNVTNFVAORLSAELL--LGQLNAEHAADV 102
DB 133 BECRPPSLI-----TRVSYFAVFGHGIRASKFAAQNHLNIRKPKGDIVSVKTV 187
QY 103 RRVLLQAFDVVERSFLSIDDALAEKASQLQPEGVPQHLPQYOKILERLKTLEI 162
DB 188 KRCLLDTFKHDEEFL-----KQASSQK-----PAWK----- 214
QY 163 SGGMAVAVLLNNKLYVANYGTNRALLCKSTVDGLQ--VTQLNVDHHTENEDELFRLSQ 220
DB 215 -DGSTATCVLAVDNLIYLANLGDRAILCRYNESQKHAALSLSKEHNPTQYERMRIQK 273
QY 221 LGLDAGKIKVGIICQBSTRIDG--YVKYGYTDIDLLSAASKPIIAEPEIHGAO-- 276
DB 274 AG--GNVRDGRVLGVLEVSRSIGDQYK-RCGVTSV-----PDIRRCQLT 315
QY 277 PLDGVTFGLVMSGLYKALEAAHGPQAOEIAAMIDTEFAK---QTSLDVAQAQAVD 332
DB 316 PND---RILLACDGLFKVFT----PEAVNFIISLEDEKIQIREGSAADARYEAACN 368
QY 333 RV 334
DB 369 RL 370

RESULT 11

US-09-935-855-20
; Sequence 20, Application US/08935855
; Patent No. 6066485
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basilio, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/935,855
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 546 amino acids
; TYPE: amino acid

; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-08-935-855-20

Query Match 4.9%; Score 125.5; DB 3; Length 546;
Best Local Similarity 20.9%; Pred. No. 0.0056;
Matches 78; Conservative 65; Mismatches 134; Indels 97; Gaps 17;

QY 24 CHLSGVGSASNRYSADGKGTSHP-----PEDSWLKRSENENFLYGVNGYDG 73
DB 164 CH-----KGPHSKSGGGTGEEPSQGLNGEAGPEDSTRETPOENGPTAKAYTFSS 216
QY 74 NRVNTVAORLSAELLGQLNAEHAFA-----DVRVLLQAFDVVERSFLSIDDA 124
DB 217 N-----SERGTEAGOVGPPIPTGEAGPSCSSASDKLPRVAKSKF--PEDSEDES-DEA 267
QY 125 LAFKASLQ--SOLPEGVPQHLPQ-----YOKILERLKTLEI-----SG 164
DB 268 EEEEESECSSEEDGYSSAEAEDEDDTEAEDEDEEEEMVPMGMEKESPGSDS 327
QY 165 GAMAVAVLLNNKLYVANYGTNRALLCKSTVDGLQVTQLNVDHHTENEDELFRLSQLGLD 224
DB 328 GTTAVVALLRQQLIVANAGDSRCVVSEAG----KALDMSYDHKPEDEVELARIKNAG-- 381
QY 225 AGKIKQVG--IICQSESTRIGDYKVKYGYTDIDLSAASKPIIAEPEIHGAO----- 277
DB 382 -GKVTMDGRVNGNLNLSRAIGDHFYK-----RNKNLPPEEQMISALPDIKVLT 428
QY 278 LDGVTGFLVMSGLYKALEAAHGPQAOEIAAMIDTEFAKQTS-----LDVAQAQAV 331
DB 429 LTDDEHFMVIACDGIMNVN-----SSQEVVDFIQISKISORDENGLRLUSSIVEELL 480

QY 332 DRVRIHSDTFASG 345
DB 481 DQC--LAPDTSGDG 492

RESULT 12

US-08-822-701-8
; Sequence 8, Application US/08822701
; Patent No. 5976853
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basilio, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE:
; APPLICATION NUMBER: US/08/822,701
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 N
; TELECOMMUNICATION INFORMATION:

```

;
; TELEPHONE: 201-487-5800
;
; TELEFAX: 201-343-1684
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Rattus
;
US-08-822-701-8

```

```

Query Match 4.8%; Score 124; DB 2; Length 306;
Best Local Similarity 19.3%; Pred. No. 0.0029;
Matches 53; Conservative 49; Mismatches 110; Indels 62; Gaps 8;

```

```

QY 65 YGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLESIDDA 124
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 FAVYDGHAGSQVAKYCCCHL-----LDHITNN 82
QY 125 LAEKASLSQSLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLLNNKLYV 180
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 QDFKSGAGAPSVENV-KNGIRTFLEIDEMRMVSEKKHGAORSSGTAVGLISPOHTYF 141
QY 181 ANVTNRALLCKSTVDGLQVTLNVHTTENEDLFRLSQLGLDAGKIKOVGIICQEST 240
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 INGDSTRGLLCNR---KVHFTQDHKPSNPLEKERIQNAG---GSVMIGRVNSGLAVS 194
QY 241 RRIQ--DYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALAA 298
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 RALGDPDYKCVHKGPTQL-----VSPEPEVHDIERSEDDQFIILACGIMDVM-- 245
QY 299 AHGPGQANOETIAMIDTEFAKQTSIDAVAGAYVD 332
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 246 -----GNEELCDFVRSRLLEVTDLEKVCNEVVD 273

```

```

RESULT 13
US-08-935-855-8
; Sequence 8, Application US/08935855
; Patent No. 6066485
; GENERAL INFORMATION:
; APPLICANT: Guthridge, Mark
; APPLICANT: Basilio, Claudio
; TITLE OF INVENTION: NOVEL GROWTH FACTOR INDUCIBLE
; TITLE OF INVENTION: SERINE/THREONINE PHOSPHATASE, FIN13
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/935,855
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 1049-1-002 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800

```

```

;
; TELEFAX: 201-343-1684
;
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; ORGANISM: Rattus
;
US-08-935-855-8

```

```

Query Match 4.8%; Score 124; DB 3; Length 306;
Best Local Similarity 19.3%; Pred. No. 0.0029;
Matches 53; Conservative 49; Mismatches 110; Indels 62; Gaps 8;

```

```

QY 65 YGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLESIDDA 124
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 56 FAVYDGHAGSQVAKYCCCHL-----LDHITNN 82
QY 125 LAEKASLSQSLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLLNNKLYV 180
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 83 QDFKSGAGAPSVENV-KNGIRTFLEIDEMRMVSEKKHGAORSSGTAVGLISPOHTYF 141
QY 181 ANVTNRALLCKSTVDGLQVTLNVHTTENEDLFRLSQLGLDAGKIKOVGIICQEST 240
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 142 INGDSTRGLLCNR---KVHFTQDHKPSNPLEKERIQNAG---GSVMIGRVNSGLAVS 194
QY 241 RRIQ--DYKVKYGYTDIDLLSAKSKPIIAEPEIHGAQPLDGVTFGLVMSEGLYKALAA 298
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 RALGDPDYKCVHKGPTQL-----VSPEPEVHDIERSEDDQFIILACGIMDVM-- 245
QY 299 AHGPGQANOETIAMIDTEFAKQTSIDAVAGAYVD 332
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 246 -----GNEELCDFVRSRLLEVTDLEKVCNEVVD 273

```

```

RESULT 14
US-09-206-646-4
; Sequence 4, Application US/09206646
; Patent No. 6436637
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; APPLICANT: Lal, Preeti G.
; APPLICANT: Corley, Neil C.
; APPLICANT: Zhang, Hong
; TITLE OF INVENTION: NEW PROTEIN PHOSPHATASE
; FILE REFERENCE: PF-0319-1 DIV
; CURRENT APPLICATION NUMBER: US/09/206,646
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ. ID NOS: 4
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: ID No. 6436637 g452526
;
US-09-206-646-4

```

```

Query Match 4.6%; Score 118; DB 4; Length 390;
Best Local Similarity 21.4%; Pred. No. 0.016;
Matches 60; Conservative 52; Mismatches 107; Indels 62; Gaps 12;

```

```

QY 60 NNCFYGVFNGYDGNRYTNFVAQRLSAELLGQLNAEHAADVRRVLLQAFDVERSFLES 119
; : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 51 DNMSFVAVYDGHAGSRVAVYCSHLLLEHI---TTNEDFRAD-----KSGSLAES-VE 100
QY 120 SIDDALAEKASLSQSLPEGVPOHLPPOYOKILERLKTLEIREISG---GAMAVVAVLLN 175

```


1	2580	100.0	504	9	US-10-158-895-2	Sequence 2, Appli
2	2580	100.0	504	12	US-10-123-427-2	Sequence 2, Appli
3	2580	100.0	517	9	US-10-158-895-11	Sequence 11, Appli
4	2577	99.9	513	9	US-10-158-895-43	Sequence 43, Appli
5	2575	99.8	504	12	US-10-123-427-6	Sequence 6, Appli
6	409	15.9	84	10	US-09-825-300-1270	Sequence 1270, Appli
7	265	14.1	70	10	US-09-864-761-34065	Sequence 34065, A
8	261	10.1	51	10	US-09-864-761-34067	Sequence 34067, A
9	205	7.9	274	10	US-09-860-351-4	Sequence 4, Appli
10	178	6.9	392	9	US-09-935-124A-2	Sequence 2, Appli
11	170.5	6.6	372	9	US-09-973-941-4	Sequence 4, Appli
12	170.5	6.6	372	10	US-09-973-963-4	Sequence 4, Appli
13	170.5	6.6	372	10	US-09-973-064-4	Sequence 4, Appli
14	170.5	6.6	372	10	US-09-973-077-4	Sequence 4, Appli
15	170.5	6.6	372	10	US-09-973-083-4	Sequence 4, Appli
16	170.5	6.6	372	10	US-09-973-964-4	Sequence 4, Appli
17	170.5	6.6	372	10	US-09-975-072-4	Sequence 4, Appli
18	170.5	6.6	372	10	US-09-972-038-4	Sequence 4, Appli
19	170.5	6.6	372	10	US-09-972-757-4	Sequence 4, Appli

```

QY 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
|
|
|
Db 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
QY 241 RRIQGYKVKYGYTDDIDLSAAKSKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
|
|
|
Db 241 RRIQGYKVKYGYTDDIDLSAAKSKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
QY 301 GGGQANOIEAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
|
|
|
Db 301 GGGQANOIEAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNRFQYPLGEMSQPTSPAPAAAGRVYVSVYSSAOSTSKTSVTLSLVMPSCQGMVNG 420
|
|
|
Db 361 LVNRFQYPLGEMSQPTSPAPAAAGRVYVSVYSSAOSTSKTSVTLSLVMPSCQGMVNG 420
QY 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
|
|
|
Db 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
|
|
|
Db 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

```

```

RESULT 2
US-10-123-427-2
; Sequence 2, Application US/10123427
; Patent No. US20020119525A1
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihito
; NISHIDA, Eisuke
; TITLE OF INVENTION: TABI PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/123,427
; FILING DATE: 17-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

```

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-123-427-2
Query Match 100.0%; Score 2580; DB 12; Length 504;
Best Local Similarity 100.0%; Pred. No. 5e-194;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRSLSEQQPSWTDDLPICLSGVGSASNSRYSADGKGTSHPEPDSWLKFRSEN 60
|
|
|
Db 1 MAQRSLSEQQPSWTDDLPICLSGVGSASNSRYSADGKGTSHPEPDSWLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRYTNFVAORLSAELLGOLNAEHAADYRVLLOAFDVERSFLS 120
|
|
|
Db 61 NCFLYGVFNQYDGNRYTNFVAORLSAELLGOLNAEHAADYRVLLOAFDVERSFLS 120
QY 121 IDDALEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGMAVAVALNNKLYV 180
|
|
|
Db 121 IDDALEKASLQSLPEGVPOHLPPOYOKILERLKTLEREISGMAVAVALNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
|
|
|
Db 181 ANVTNRALLCKSTVDGLQVTLQNVNHTTENEDELFRLSQGLDAGKIKQVGIICQEST 240
QY 241 RRIQGYKVKYGYTDDIDLSAAKSKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
|
|
|
Db 241 RRIQGYKVKYGYTDDIDLSAAKSKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
QY 301 GGGQANOIEAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
|
|
|
Db 301 GGGQANOIEAAMIDTEFAKQTSIDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNRFQYPLGEMSQPTSPAPAAAGRVYVSVYSSAOSTSKTSVTLSLVMPSCQGMVNG 420
|
|
|
Db 361 LVNRFQYPLGEMSQPTSPAPAAAGRVYVSVYSSAOSTSKTSVTLSLVMPSCQGMVNG 420
QY 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
|
|
|
Db 421 AHSASTLDEATPTLTNOSTPTLTQSTNTHHTOSSSSSDGGLFRSPAHSLPGEEDRVER 480
QY 481 YVDFAEFYRLMSVDHGEGSVVTAP 504
|
|
|
Db 481 YVDFAEFYRLMSVDHGEGSVVTAP 504

```

```

RESULT 3
US-10-158-895-11
; Sequence 11, Application US/10158895
; Patent No. US20020155624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053469/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; PRIOR FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-11
Query Match 100.0%; Score 2580; DB 9; Length 517;
Best Local Similarity 100.0%; Pred. No. 5.2e-194;
Matches 504; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MAQRRLSLOEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 1 MAQRRLSLOEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
QY 61 NCFLYGVFNQYDGNRVNFAQRLSABELLGQLNAEHAADVRVLLQAFDVVERSFLS 120
Db 61 NCFLYGVFNQYDGNRVNFAQRLSABELLGQLNAEHAADVRVLLQAFDVVERSFLS 120
QY 121 IDALAEKASIQSOLPEGVPOHOLPPQYQKILERLKTLEREISGGMAMVAVVLLNNKLYV 180
Db 121 IDALAEKASIQSOLPEGVPOHOLPPQYQKILERLKTLEREISGGMAMVAVVLLNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFRLSQLGLDAGKI KQVGIICQEST 240
Db 181 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFRLSQLGLDAGKI KQVGIICQEST 240
QY 241 RRIQDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
Db 241 RRIQDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
QY 301 GPGQANQEI AAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 301 GPGQANQEI AAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYPVSPYSSAQSTSKTSVTLSLVMPSSQGMVNG 420
Db 361 LVRNFGYPLGEMSOPTSPAPAGGRVYPVSPYSSAQSTSKTSVTLSLVMPSSQGMVNG 420
QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
QY 481 YVDFAEFYRLMSVDHGQSVVVTAP 504
Db 481 YVDFAEFYRLMSVDHGQSVVVTAP 504

RESULT 4
US-10-158-895-43
; Sequence 43, Application US/10158895
; Patent No. US20020153624A1
; GENERAL INFORMATION:
; APPLICANT: ONO, KOICHIRO
; APPLICANT: OHTOMO, TOSHIHIKO
; APPLICANT: TSUCHIYA, MASAYUKI
; TITLE OF INVENTION: METHOD OF SCREENING TGF-BETA INHIBITORY SUBSTANCES
; FILE REFERENCE: 053466/0278
; CURRENT APPLICATION NUMBER: US/10/158,895
; CURRENT FILING DATE: 2002-06-03
; PRIOR APPLICATION NUMBER: US/09/529,279
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: PCT/JP98/04796
; PRIOR FILING DATE: 1998-10-22
; PRIOR APPLICATION NUMBER: JP 9/290188
; PRIOR FILING DATE: 1997-10-22
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 43
; LENGTH: 513
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-158-895-43

Query Match 99.9%; Score 2577; DB 9; Length 513;
Best Local Similarity 99.8%; Pred. No. 8.8e-194;
Matches 503; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MAQRRLSLOEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 60
Db 10 MAQRRLSLOEQPSWTDLPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSEN 69
QY 61 NCFLYGVFNQYDGNRVNFAQRLSABELLGQLNAEHAADVRVLLQAFDVVERSFLS 120

Db 70 NCFLYGVFNQYDGNRVNFAQRLSABELLGQLNAEHAADVRVLLQAFDVVERSFLS 129
QY 121 IDALAEKASIQSOLPEGVPOHOLPPQYQKILERLKTLEREISGGMAMVAVVLLNNKLYV 180
Db 130 IDALAEKASIQSOLPEGVPOHOLPPQYQKILERLKTLEREISGGMAMVAVVLLNNKLYV 189
QY 181 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFRLSQLGLDAGKI KQVGIICQEST 240
Db 190 ANVTNRALLCKSTVDGLQVTLQNVHTTENEDLFRLSQLGLDAGKI KQVGIICQEST 249
QY 241 RRIQDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 300
Db 250 RRIQDYKVKYGYTDIDLLSAASKPIIAEPEIHGAQPLDGVTFVLMSEGLYKALEAAH 309
QY 301 GPGQANQEI AAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 360
Db 310 GPGQANQEI AAMIDTEFAKOTSLDAVAQAVVDRVKRIHSDTFASGGERARFCPRHEDMTL 369
QY 361 LVRNFGYPLGEMSOPTSPAPAGGRVYPVSPYSSAQSTSKTSVTLSLVMPSSQGMVNG 420
Db 370 LVRNFGYPLGEMSOPTSPAPAGGRVYPVSPYSSAQSTSKTSVTLSLVMPSSQGMVNG 429
QY 421 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 480
Db 430 AHSASTLDEATPTLTNOSPPTLTQSTNTHQTSSSSSDGGLFRSRPAHSLPPGEDGRVEP 489
QY 481 YVDFAEFYRLMSVDHGQSVVVTAP 504
Db 490 YVDFAEFYRLMSVDHGQSVVVTAP 513

RESULT 5
US-10-123-427-6
; Sequence 6, Application US/10123427
; Patent No. US20020119525A1
; GENERAL INFORMATION:
; APPLICANT: MATSUMOTO, Kunihiro
; APPLICANT: NISHIDA, Eisuke
; TITLE OF INVENTION: TAB1 PROTEIN AND DNA CODING THEREFOR
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/123,427
; FILING DATE: 17-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/406,854
; FILING DATE: <Unknown>
; APPLICATION NUMBER: US/08/752,891
; FILING DATE: 20-NOV-1996
; APPLICATION NUMBER: JP 8-300856
; FILING DATE: 28-OCT-1996
; APPLICATION NUMBER: JP 8-126282
; FILING DATE: 24-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 17981/111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399

```
TELEX: 904136
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 504 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-123-427-6

Query Match          99.8%; Score 2575; DB 12; Length 504;
Best Local Similarity 99.8%; Pred. No. 1.2e-193;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAAQRSLIQSEQSPWTDLPLCHLSGVGSASNSYSADGKGTSHPEDSWLKFRSEN 60
DB 1 MAAQRSLIQSEQSPWTDLPLCHLSGVGSASNSYSADGKGTSHPEDRWLKFRRSEN 60
QY 61 NCFLYGVFNQGNRVNTNVAORLSAELLGQLNMAHADVRVYLQAFDVERSPLFS 120
DB 61 NCFLYGVFNQGNRVNTNVAORLSAELLGQLNMAHADVRVYLQAFDVERSPLFS 120
QY 121 IDDLALAEKASLSQSLPEGVPOHQLPPQYOKILERLKTTEREISGGAMAVAVLLNNKLYV 180
DB 121 IDDLALAEKASLSQSLPEGVPOHQLPPQYOKILERLKTTEREISGGAMAVAVLLNNKLYV 180
QY 181 ANVTNRALLCKSTVDGLQVTLQVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIGQEST 240
DB 181 ANVTNRALLCKSTVDGLQVTLQVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIGQEST 240
QY 241 RRTGDKVKKVYGTIDILSAASKPIIARPEIHGAQPLDGVGVFLVMSGGYKALAEAAH 300
DB 241 RRTGDKVKKVYGTIDILSAASKPIIARPEIHGAQPLDGVGVFLVMSGGYKALAEAAH 300
QY 301 GPGQANQEIAMIDTEFAKQTSIDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
DB 301 GPGQANQEIAMIDTEFAKQTSIDAVAQAQVVDVVKRIHSDTFASGGERARFCPRHEDMTL 360
QY 361 LVNRFQYPLGMSQPTPSPAPAGGRVYVSVYSSAQSTSTKTSVTLSLWMSQGMVNG 420
DB 361 LVNRFQYPLGMSQPTPSPAPAGGRVYVSVYSSAQSTSTKTSVTLSLWMSQGMVNG 420
QY 421 AHSASTLDEATPTLTNOSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSLPGEDEGRVPE 480
DB 421 AHSASTLDEATPTLTNOSPTLTLOSTNHTQSSSSSDGGLFRSRPAHSLPGEDEGRVPE 480
QY 481 YVDFAEFYRLMSYDHGEQSVVTAP 504
DB 481 YVDFAEFYRLMSYDHGEQSVVTAP 504

RESULT 6
US-09-925-300-1270
; Sequence 1270, Application US/09925300
; Patent No. US20020351681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatencIn Ver. 2.0
; SEQ ID NO 1270
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1270

Query Match          15.9%; Score 409; DB 10; Length 84;
Best Local Similarity 97.5%; Pred. No. 1.5e-25;
Matches 78; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 425 STUDEATPTLTNOSPPTLTLOSTNHTQSSSSSDGGLFRSRPAHSLPGEDEGRVPEYDF 484
DB 5 ATLDEATPTLTNOSPPTLTLOSTNHTQSSSSSDGGLFRSRPAHSLPGEDEGRVPEYDF 64
QY 485 AEFYRLMSYDHGEQSVVTAP 504
DB 65 AEFYRLMSYDHGEQSVVTAP 84

RESULT 7
US-09-864-761-34065
; Sequence 34065, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34065
; LENGTH: 70
; TYPE: PRT
```



```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.77
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.4
; OTHER INFORMATION: EST_HUMAN HIT: BE746542.1, EVALUE 5.00e-35
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 4.00e-36
US-09-864-761-34065

Query Match 14.1%; Score 365; DB 10; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e-22;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 312 MIDTEFAKQTSLDVAQAQVDRVKRIHSDTTFASGGERARFCPRHEDMTLLVRNFGYPLGE 371
Db 1 MIDTEFAKQTSLDVAQAQVDRVKRIHSDTTFASGGERARFCPRHEDMTLLVRNFGYPLGE 60

Qy 372 MSQPTSPAP 381
Db 61 MSQPTSPAP 70

RESULT 8
US-09-864-761-34067
; Sequence 34067, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aesomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34067
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO Z83845.14
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 10
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.9
; OTHER INFORMATION: SWISSPROT HIT: Q15750, EVALUE 7.00e-24
; OTHER INFORMATION: EST_HUMAN HIT: AL118967.1, EVALUE 9.00e-23
US-09-864-761-34067

Query Match 10.1%; Score 261; DB 10; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.7e-14;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 SENNCFLYGVNGYDGNRVTFNFAQRLSABLLGQLNAEHAEDVRRVLLQ 108
Db 1 SENNCFLYGVNGYDGNRVTFNFAQRLSABLLGQLNAEHAEDVRRVLLQ 51

RESULT 9
US-09-860-351-4
; Sequence 4, Application US/09860351
; Patent No. US20020077463A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 16105, A NOVEL PROTEIN HUMAN PHOSPHATASE
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: 38155-20013.00
; CURRENT APPLICATION NUMBER: US/09/860,351
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/205,260
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid
US-09-860-351-4

Query Match 7.9%; Score 205; DB 10; Length 274;
Best Local Similarity 24.3%; Pred. No. 7.2e-09;
Matches 74; Conservative 53; Mismatches 103; Indels 74; Gaps 12;

Qy 34 NRSYSADGKGTEHPPEDSWLKFRSENNCFLYGVNGYDGNRVTFNFAQRLSABLLGQL 93
Db 13 NLNSSSGK-----DSW-----SFFAVDFHGGGQAAYAKGHLHKTILAE- 54

Qy 94 NAEHAEDVRRVLLQAF-DVVERSEFLESIDDALEAKSLQSLPEGVPHQUPPOYQKIL 152
Db 55 -KSFPEGDPMWKLSDELDALKESFLEADTDELRSAASAA-----NKVL 99
```


; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: US 60/304,775
; PRIOR FILING DATE: 2001-07-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-963-4

Query Match 6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDPAQLTDEVLYFAVYDGHGGAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLIESIDDAEAKASLOSQ 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHLPPOYQKILERLKTLEIREISGGAMAVAVLLNN-KLYVANVTNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDGRAILCRKG-- 217
Qy 197 GLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC-----GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKRIKKCG-----GFVAVNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308
Qy 299 -----AHGPGQANQETAAAMIDTEFAKQTSLDAVAQAVV-----DRVKRIHSD 340
Db 309 ICDFVNCQCHDPNEAAHAV-----TEQAIOYGTEDNSTAVVVPFGAWGKYKNSINFSPSR 363
Qy 341 TFASGGERA 349
Db 364 SFASSGRWA 372

RESULT 13
US-09-973-064-4
; Sequence 4, Application US/09973064
; Patent No. US20020106773A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,064
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-064-4

Query Match 6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDPAQLTDEVLYFAVYDGHGGAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLIESIDDAEAKASLOSQ 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHLPPOYQKILERLKTLEIREISGGAMAVAVLLNN-KLYVANVTNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDGRAILCRKG-- 217
Qy 197 GLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC-----GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKRIKKCG-----GFVAVNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308

Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDPAQLTDEVLYFAVYDGHGGAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLIESIDDAEAKASLOSQ 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHLPPOYQKILERLKTLEIREISGGAMAVAVLLNN-KLYVANVTNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDGRAILCRKG-- 217
Qy 197 GLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC-----GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKRIKKCG-----GFVAVNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308
Qy 299 -----AHGPGQANQETAAAMIDTEFAKQTSLDAVAQAVV-----DRVKRIHSD 340
Db 309 ICDFVNCQCHDPNEAAHAV-----TEQAIOYGTEDNSTAVVVPFGAWGKYKNSINFSPSR 363
Qy 341 TFASGGERA 349
Db 364 SFASSGRWA 372

RESULT 14
US-09-973-077-4
; Sequence 4, Application US/09973077
; Patent No. US20020114799A1
; GENERAL INFORMATION:
; APPLICANT: Roch, Jean-Marc
; APPLICANT: Bartel, Paul L.
; APPLICANT: Heichman, Karen
; TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
; FILE REFERENCE: Protein Interactions in ND
; CURRENT APPLICATION NUMBER: US/09/973,077
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: US 60/240,790
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 4
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-973-077-4

Query Match 6.6%; Score 170.5; DB 10; Length 372;
Best Local Similarity 24.7%; Pred. No. 5.5e-06;
Matches 91; Conservative 50; Mismatches 105; Indels 123; Gaps 19;

Qy 21 LPLCHLSGVGSASNRYSADGKGTESHPPEDSWLKFRSENNCFLYGVFNGYDGNRVTFV 80
Db 87 IPKISLENVGCSAQ-----IGKRKEN---EDRFDPAQLTDEVLYFAVYDGHGGAADFC 138
Qy 81 ---AQRLSAELLGQLNAEHAADVRRVLLQAFDVVERSFLIESIDDAEAKASLOSQ 137
Db 139 HTHMEKICMDLLPKEKNLE-----TLLTLAFLE-IDKAFSSHARLSAD--- 180
Qy 138 GVPQHLPPOYQKILERLKTLEIREISGGAMAVAVLLNN-KLYVANVTNRALLCKSTVD 196
Db 181 -----ATL---LTSGTTATVALLRDGIELVVASVGDGRAILCRKG-- 217
Qy 197 GLQVTLNVNDHTTENEDELFRLSQLGLDAGKIKQVGIIIC-----GQ-----ESTRIGD 245
Db 218 --KPMKLTIDHTPERKDEKRIKKCG-----GFVAVNSLQOPHVNGLAMTRSIG- 265
Qy 246 YKVKYGYTIDLLSAKSKPIIAEPE-----IHGAQPLDGVGTGFLVLMSEGLYKALEA-- 298
Db 266 -----DLDL-----KTSGVIAEPETKRIKLHHAD-----DSFLVLTDTGDNFMVNSQE 308

```

QY 229 -----AHGGGANOELAMIDTEPAKOTSLDAVAOV-----DRYKRLSD 340
      : : : : :
Db 309 ICDPYNQCHDNEEAHAY-----TEQAIQYTEDNSTVAVVPCGAWGKYNNSEINFSPR 363
      : : : : :
QY 341 TPASGGERA 349
      : : : : :
Db 364 SPASGGRWA 372

```

RESULT 15

```

US-09-973-063-4
: Sequence 4, Application US/09973063
: Patent No. US20020115119A1
: GENERAL INFORMATION:
: APPLICANT: Koch, Jean-Marc
: APPLICANT: Bartel, Paul L.
: APPLICANT: Hetchman, Karen
: TITLE OF INVENTION: Protein-Protein Interactions in Neurodegenerative
: TITLE OF INVENTION: Diseases
: FILE REFERENCE: Protein Interactions in ND
: CURRENT APPLICATION NUMBER: US/09/973,063
: CURRENT FILING DATE: 2001-10-10
: PRIOR APPLICATION NUMBER: US/60/240,790
: PRIOR FILING DATE: 2000-10-17
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 4
: LENGTH: 372
: TYPE: PRN
: ORGANISM: Homo sapiens
US-09-973-063-4

```

Query Match	6.6%	Score 170.5;	DB 10;	Length 372;
Best Local Similarity	24.7%;	Pred. No. 5.5e+06;		
Matches 91; Conservative	50;	Mismatches 105;	Indels 123;	Gaps 19;

Search completed: December 9, 2002, 23:04:26
Job time : 101.986 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:58:34 ; Search time 3162.68 Seconds
(without alignments)
4637.778 Million cell updates/sec

Title: US-09-830-144-4
Perfect score: 2580
Sequence: 1 MAQRRLSQEQPQSWTDD.....AEFYRLMSVDHGQSVVTAP 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODEL=frame_p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09830144/runat_04122002_141353_2252/app_query.fasta_1.1422
-DB=GenEmbl -QFMT=fastcap -SUFFIX=p2n.rge -MINWATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG 1.1 5173 @runat 04122002_141353_2252 -NCPU=6 -ICPU=3
-NO_XUPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sv.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*

- 29: em.vi.*
- 30: em.htg_hum.*
- 31: em.htg_inv.*
- 32: em.htg_other.*
- 33: em.htg_mus.*
- 34: em.htg_pln.*
- 35: em.htg_rtd.*
- 36: em.htg_nam.*
- 37: em.htg_vrt.*
- 38: em.sv.*
- 39: em.htgo_hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2580	100.0	1515	6	E31041	E31041 Method for
2	2580	100.0	1560	6	AR058299	AR058299 Sequence
3	2580	100.0	1560	6	AR088273	AR088273 Sequence
4	2580	100.0	1560	6	AR116881	AR116881 Sequence
5	2580	100.0	1560	6	E14752	E14752 Human mRNA
6	2580	100.0	3096	9	HSU49928	U49928 Homo sapien
7	2575	99.8	1560	6	AR058302	AR058302 Sequence
8	2575	99.8	1560	6	AR088276	AR088276 Sequence
9	2575	99.8	1560	6	AR116884	AR116884 Sequence
10	2512	97.4	2944	10	BC027054	BC027054 Mus muscu
11	1924	74.6	1926	5	XLU92031	U92031 Xenopus lae
12	547.5	21.2	69660	9	HS407F17	Z83845 Human DNA s
13	457	17.7	176685	2	AC127784	AC127784 Rattus no
14	414	16.0	176685	2	AC127784	AC127784 Rattus no
15	387	15.0	141498	2	AC127924	AC127924 Rattus no
16	369	14.3	141498	2	AC127924	AC127924 Rattus no
17	258.5	10.0	1161	3	AF145375	AF145375 Caenorhab
18	208	8.1	1479	8	AB083482	AB083482 Mesembrya
19	205.5	8.0	967	8	AY133761	AY133761 Arabidops
20	205.5	8.0	1453	8	AY080658	AY080658 Arabidops
21	203.5	7.9	1824	6	AX319337	AX319337 Sequence
22	202	7.8	1851	8	AY086406	AY086406 Arabidops
23	198	7.7	1463	8	AB079671	AB079671 Arabidops
24	198	7.7	1522	8	AY050873	AY050873 Arabidops
25	196.5	7.6	1099	8	AY031209	AY031209 Arabidops
26	194.5	7.5	849	8	AF370608	AF370608 Arabidops
27	194	7.5	31332	8	AP005244	AP005244 Oryza sat
28	193.5	7.5	852	8	AY133563	AY133563 Arabidops
29	193.5	7.5	96221	8	AC051631	AC051631 Arabidops
30	193.5	7.5	109367	8	AC023673	AC023673 Genomic s
31	192.5	7.5	1194	8	AY048244	AY048244 Arabidops
32	192	7.4	1345	8	AY080735	AY080735 Arabidops
33	190.5	7.4	1422	8	AY085196	AY085196 Arabidops
34	190.5	7.4	1531	8	AY057611	AY057611 Arabidops
35	190.5	7.4	1902	8	AY062454	AY062454 Arabidops
36	189	7.3	877	8	AY133737	AY133737 Arabidops
37	188.5	7.3	1074	8	AY113024	AY113024 Arabidops
38	188.5	7.3	1479	8	AY087522	AY087522 Arabidops
39	187.5	7.3	1259	8	AY084874	AY084874 Arabidops
40	187	7.2	2204	8	AY065026	AY065026 Arabidops
41	186.5	7.2	1689	8	AF092587	AF092587 Arabidops
42	186	7.2	1248	8	AF092432	AF092432 Lotus jap
43	186	7.2	1767	8	AF499718	AF499718 Thellungi
44	185	7.2	1966	8	AY045819	AY045819 Arabidops
45	184.5	7.2	1263	8	AY133656	AY133656 Arabidops

ALIGNMENTS

RESULT 1

E31041
LOCUS E31041 1515 bp DNA linear PAT 18-JUN-2001
DEFINITION Method for screening substance inhibiting binding to XIAP.
ACCESSION E31041
VERSION E31041.1 GI:13017306
KEYWORDS JP 199326328-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1515)
AUTHORS Kunhiro,M.
TITLE Method for screening substance inhibiting binding to XIAP
JOURNAL Parent: JP 199326328-A 1 26-NOV-1999;
KUNHIRO MATSUMOTO
COMMENT OS Unidentified
PN JP 199326328-A/1
PD 26-NOV-1999
PF 13-MAY-1998 JP 1998130378
PR
PI KUNHIRO MATSUMOTO
PC GOIN33/536, GOIN33/536//C12N15/09, C12P21/08, A61K37/02,
A61K37/02, PC, A61K37/02, C12N15/00
PC A61K39/395,
PC A61K45/00, A61K45/00, A61K45/00, C07K7/06, C07K7/08, C07K14/47, PC
GOIN33/536,
PC GOIN33/536, GOIN33/536//C12N15/09, C12P21/08, A61K37/02,
A61K37/02, PC, A61K37/02, C12N15/00
CC Strandedness: Double;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS 1..1512.
FEATURES
source 1..1515 Location/Qualifiers
BASE COUNT 323 a 457 c 463 g 272 t
ORIGIN
Alignment Scores:
Pred. No.: 3,85e-151 Length: 1515
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-4 (1-504) x E31041 (1-1515)
QY 1 MetAlaAlaGlnArgSerLeuLeuGlnInserGluGlnInProSerTyrThrAsp 20
DB 1 ATGGCGGCGCAGAGAGAGAGCTTCTGCAAGTAGAGCGCAGCCAGCTGAGACAGATGAC 60
QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 61 CTGCTCTCTGCGACCTCTCTGGGGTGGCTCAGCCTCCACCGCAGCTACTCTGCTGAT 120
QY 41 GlyLysGlyThrGlnSerHisProProGluAspSerTyrPheArgSerGluAsn 60
DB 121 GGCAAGGGCAGCTGAGAGCCAGCCGCGCAGAGACAGCTGGCTCAGAGTTCCAGAGTGAGAAC 180
QY 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 181 AACTGCTTCTCTGTATGGGGCTCTTCAACGGCTATGATGCAACCGATGACCACTTCTGTG 240
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlnGlyGlnLeuAsnAlaGluHisAlaGluAla 100
DB 241 GCCCAGCGGCTGTCCGCGAGAGCTCTGCTGGCCAGCTGAAATGCCAGCAGCAGAGGCC 300
QY 101 AspValArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
DB 301 GATGTGGCGGTGTGTCTGCTCAGGCCCTTCATGTGTGGAGAGAGAGCTTCTGTGAGTCC 360
QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnInserGluLeuProGluGlyValPro 140

DB 361 ATTACACCGCTTGGCTGGAAGGCGCAAGCCTCAGTCGCAATTCGACAGGAGTCCCT 420
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuSerThrLeuGlnArg 160
DB 421 CAGCACCACTGCTCTCTCTGATTCAGATCAGAAAGATCTTGAGACTCAAGCGTTAGAGAG 480
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 481 GAAATTCGGAGAGGGCCATGCGCGTTGTGGCGCTCTTCTCAACAACAAGCTTACGTC 540
QY 181 AlaAsnValGlyThrAsnAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
DB 541 GCCAATGTCGTACCAACCGTCGACTTTATGCAAAATTCACAGTGGATGGTGACGTG 600
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnLysGluAspGluLeuPheArgLeuSerGln 220
DB 601 ACACAGCTGAACGTGAGCCACACACAGAGAGAGATGAGCTTCCGCTCTTTCGACG 660
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCySerGlyGlnInserThr 240
DB 661 CTGGGCTTGGATGCTGGAAAGATCAAGCAGGTGGGATCATCTGTGGCAGGAGACAC 720
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAla 260
DB 721 CGGGGATCGGGGATTACAGGTTAAATGCTTACACGACATTCACCTTCCAGCGCT 780
QY 261 AlaSerLysProIleIleAlaGluProGluIleHisGlyValAlaGlnProLeuAspGly 280
DB 781 GCCAAGTCCAAACCATCATCGAGAGCCAGAAATCATGGGCGACAGCCGCTGGATGG 840
QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis 300
DB 841 GTGACGGGCTTCTGTGTGCTGATGCGAGGGGTGTGTCAGAGCCCTTAGAGCAGCCAT 900
QY 301 GlyProGlyGlnLysGlnGlnIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
DB 901 GGGCTGTGGGAGGCGCAACGAGAGATTGCTGCATGATTAACATGATTCGCAAGCAG 960
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
DB 961 ACTTCCTGGACCGCATATGGCCAGAGCGTCGTGAGACGGGTAAAGCGCATTCACAGCGAC 1020
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCySProArgHisGluAspMetThrLeu 360
DB 1021 ACCTTGCCAGTGTGGGAGCGTGCCAGGTTCTGCCCCGCGCAGACAGATGACCTGTG 1080
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
DB 1081 CTATGAGGAACCTTGTGGCTAACCGGCTGGGCGAAATGAGCCAGACCCAGCGCCAGCC 1140
QY 381 ProAlaAlaGlyValArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
DB 1141 CCACTGCGAGAGAGAGATGATTAACCTGTGTGTGTCATATCTCAGAGCCAGAGAC 1200
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly 420
DB 1201 AGCAAGACCAAGCTGACCTCTCTGTCATGCCCTCCAGAGGCGCAGATGGTCAACGGG 1260
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1261 GCTCAAGTGTCTTCAACCTTGAGCAAGCCACCCACCTTACCAACCAAGCCCGGACC 1320
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnInserSerSerSerSerAspGlyGly 460
DB 1321 TTAACCTTCAGTCCACCAACACGACAGCAGACAGTCCAGCTTCAGCGAGAGCC 1380
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
DB 1381 CTCTTCCGCTCCCGGCGCGCCACCTGCTCCGCTGGCGAGAGAGCGTGTGTAAGCC 1440
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrPheSerValAspHisGlyGlnInserVal 500

Db 1441 TATGTGGACTTTGCTGAGTATTTTACCGCTCTGGAGCGTGGACCATGGCGAGCAGCGTG 1500

Qy 501 ValThrAlaPro 504

Db 1501 GTGACAGCACCG 1512

RESULT 2

AR058299

LOCUS AR058299 1560 bp DNA linear PAT 29-SEP-1999

DEFINITION Sequence 1 from patent US 5837819.

ACCESSION AR058299

VERSION AR058299.1 GI:5983876

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE Unclassified.

AUTHORS Matsumoto,K. and Nishida,E.

TITLE TAB1 protein

JOURNAL Patent: US 5837819-A 1 17-NOV-1999;

FEATURES

source Location/Qualifiers

1..1560

/organism="unknown"

BASE COUNT 332 a 469 c 480 g 279 t

ORIGIN

Alignment Scores:

Pred. No.: 3,99e-151 Length: 1560

Score: 2580.00 Matches: 504

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-830-144-4 (1-504) x AR058299 (1-1560)

Qy 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20

Db 30 ATGGCGCGCAGAGGAGGAGCTTGCTGCAGAGTGAGCAGCAGCCAAAGCTGGACAGATGAC 89

Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40

Db 90 CTGCCTCTCTGCCACCTCTCTGGGTTGGCTCAGCCTCCAACCGCAGCTACTCTGTCTGAT 149

Qy 41 GlyValGlyThrGlnSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60

Db 150 GGCAAGGGCAGTGGAGGCCACCCGCCAGAGCAGCTGGCTCAAGTTCAGGAGTGAGAAC 209

Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80

Db 210 AACTGCTTCTGTATGGGGTCTTCAAGGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269

Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100

Db 270 GCCCAGCGGCTGTCCGACAGCTCTCTGGGCAGCTGAATGCCAGCAGCCGCGAGGCC 329

Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120

Db 330 GATGTGGCGGTGTGCTGCTGAGGCCCTTCGATGTGGTGGAGGAGCTTCCTGGAGTCC 389

Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140

Db 390 ATTGACGACGCTTGGCTGAGAGGCCAAGCCTCCAGTCGCAATTGCCAGAGGAGTCCCT 449

Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160

Db 450 CAGCACCAAGCTGCCTCTCAGTATCAGAAGATCCTTGAGAGACTCAAGACGTTAGAGAG 509

Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnLeuLysLeuTyrVal 180

Db 510 GAAATTTGGGAGGGGCATGGCGCTGTGGCGGTCTTCTCAACACAAAGCTCTACGTC 569

Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200

Db 570 GCCAATGTCGGTCAAAACCGTGCACCTTTTATGCAAAATCGACAGTGGATGGTGTGCAGGTG 629

Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220

Db 630 ACACAGCTGAACGTGGACCAACACACAGAAACAGAGATGAGCTCTTCCGTCTTTTGCAG 689

Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThr 240

Db 690 CTGGCTTGGATGCTGGAAGATCAAGCAGCTGGGATCATCTGTGGCAGGAGAGCACC 749

Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260

Db 750 CGCGGATCGGGATTACAAGGTTAAATATGGCTACGACCATTTGACCTTCTCAGCGCT 809

Qy 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280

Db 810 GCCAAGTCCAAACCAATCATCGCAGACCCAGAAATCCATGGGGCAGCCGCTGGATGGG 869

Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300

Db 870 GTGACGGCTCTTGGTGCTGATGTCGAGGGGTTGTACAAGGCCCTAGAGGCGACCCAT 929

Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320

Db 930 GGGCTGGCAGGCCAAACAGGAGATTGCTCGCATGATTGACACTGAGTTTGCACAAGCAG 989

Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340

Db 990 ACCTCCCTGGACGAGTGGCCAGGCCGCTGCTGGACCGGGTGAAGCGCATCCACAGCGAC 1049

Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360

Db 1050 ACCTTCGCCAGTGTGGGAGCGTCCAGGTTCTGCCCGCCGACAGGACATGACCCCTG 1109

Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380

Db 1110 CTAGTGAGGAATTTTGGCTACCCGCTGGCGAAATGAGCCAGGCCACACCGAGCCAGCC 1169

Qy 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400

Db 1170 CCAGTGCAGAGGAGCAGGTGATCCCTGTGTGTGCCTACTCCAGGCCAGAGCACC 1229

Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420

Db 1230 AGCAAGACCGGTGACCTCTCCCTTGTCTATGCTCCCTCCAGGCCAGATGTTCAACGGG 1289

Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440

Db 1290 GCTCACAGTGTCTCACCTCGAGCAAGCCACCCCTCACCAACCAAGCCCGAC 1349

Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSer 460

Db 1350 TTAAACCTTGCAGTCCACCAACGACGACGAGCAGCAGCTCCAGCTCTGACGAGGC 1409

Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480

Db 1410 CTCCTCCCTCCCGCCCGCCACTCCGCTCCCGCTGGCAGGAGCGTCTGTTGAGGCC 1469

Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500

Db 1470 TATGTGGACTTGTCTGAGTTTTTACCGCTCTGGAGCGTGGACCATGGCGCAGAGCGTG 1529

Qy 501 ValThrAlaPro 504

Db 1530 GTGACAGCACCG 1541

RESULT 3

AR088273

LOCUS AR088273 1560 bp DNA linear PAT 07-SEP-2000

DEFINITION Sequence 1 from patent US 5989862.

ACCESSION AR088273

VERSION AR088273.1 GI:10015036

KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1560)
AUTHORS Matsumoto, K. and Nishida, E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 1 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 332 a 469 c 480 g 279 t
ORIGIN
Alignment Scores:
Pred. No.: 3,99e-151 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-830-144-4 (1-504) X AR088273 (1-1560)
QY 1 MetAlAlAGlNArgArSerLeuLeuGlnGlnGlnProSerTrpThrAsp 20
DB 30 ATGGCGGCGCAGAGGAGGAGCTTGCGAGAGTGAACAGCAGCAGCAGATGAC 89
QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
DB 90 CTCCCTCTGCGACCTCTCTGGGGTGGGCTCAGCCTCCACCGACCTACTCTCTGAT 149
QY 41 GlyVSGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
DB 150 GCGAGGGGACATGAGAGCCAGCCGCGCAGAGCAGCTGCTCAAGTTCAAGAGTGAAC 209
QY 61 AsnGlyPheLeuTyrglyValPheAsnGlyTrpAspGlyAsnArgValThrAsnPheVal 80
DB 210 AACTGCTTCTGATGGGCTTTCACAGGCTATGATGGCACCGAGTACCAACTTCGTG 269
QY 81 AlAGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGluAla 100
DB 270 GCCGACGGGCTGCTCCGACAGCTCTCGTGGGCGACGCTGAAGCCGAGACCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
DB 330 GATGTGCGGCGTGTGCTGCTGACAGCCTTCATGTGTGAGAGAGAGCTTCTCGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 390 ATTGACGAGCGCTTGCTGAGAGGCAAGCCTCCAGTCCGCAATTCGCAAGGAGATCCCT 449
QY 141 GlnHisGlnLeuProProGlnTrpGlnLysIleLeuGluArgLeuLysThrLeuGlnArg 160
DB 450 CAGCACCGACCTGCTCTCAGATCAGAAAGATCCTTGAGAGCTCAAGACGTTAGAGAGG 509
QY 161 GluIleSerGlyValAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrglyVal 180
DB 510 GAAATTTTCGGAGGGGCCATGTGGCGGTCTTCCTCAACAACAAGCTTACGCTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuGlyLysSerThrValAspGlyLeuGlnVal 200
DB 570 GCCAATGTCGTAACAACCGTCACTTTTATGCANAATGACAGTGGATGGGTGAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnLysGlnLysAspGluLeuPheArgLeuSerGln 220
DB 630 ACACAGCTGAACCTGACACACACACAGAGAAAGAGATGAGCTTCCGCTTTCCGAG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuGlyGlnGlnSerThr 240
DB 690 CTGGGCTTGAGATGCTGGAAGATCAAGCAGGTGGGATCATCTGTGGGCGAGAGGACCC 749
QY 241 ArgArgIleGlyAspTyrglyValLysTyrglyTyrrThrAspIleAspLeuLeuSerAla 260

DB 750 CGCGGATCGGGATTCAGAGTTAAATATGCTACAGCAGCATTCCTTCAGGGCT 809
QY 261 AlAlaSerLysProIleIleAlaGluProGluIleHisGlyValGlnProLeuAspGly 280
DB 810 GCCAATGCCAAACCAATCATCGCAGACCCAGAAATCCATGGGCGACGCGCTGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrglyValLeuGlnAlaHis 300
DB 870 GTGACGGGCTTCTTGCTGTGATGTGGAAGGGTTGTACAAGCCCTTACAGGAGGCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
DB 930 GGGCGTGGCAGGCCAACAGAGATTGCTGCATGATTGACATGATTGGCCAGACAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
DB 990 ACCTCCCTGAGACGAGTGGCCCAAGCCGCTGTGACCGGGTGAAGCGCATCCACCGAG 1049
QY 341 ThrPheAlaSerGlyGlyAlaArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
DB 1050 ACTTCGCCAGTGGTGGGAGCGGCGAGGTTCTGCCCCCGGACAGAGACATGACCTG 1109
QY 361 LeuValArgAsnPheGlyTyrrProLeuGlyGluMetSerGlnProThrProSerProAla 380
DB 1110 CTAGTGAGAACTTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1169
QY 381 ProAlaAlaGlyArgValTyrrProValSerValProTyrrSerSerAlaGlnSerThr 400
DB 1170 CCAAGTCAAGAGAGACAGATGATCCCTGTGTCTGTGCCATCTCCAGCGCCAGAGACCC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1230 AGCAAGACACAGCGTACCTCTCTGTGTCATGCTCCACAGGCGCAATGGTCAACGGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1290 GCTCACAGTCTTCCACCTCGGAGAGGACACCCACCTCAACCAACCAAGCCGAGCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
DB 1350 TTAAACCTGACATCCACCAACAGCAGCAGCAGAGAGACGCTTCAAGCGAGGC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
DB 1410 CTCTTCGCTCCCGGCGCCGCTCGCTCCGCTCGGAGAGACGCTGCTGTGAGGCC 1469
QY 481 TyrValAspPheAlaGluPheTyrrArgLeuTrpSerValAspHisGlyGlnGlnSerVal 500
DB 1470 TATGTGACTTGTGATGATTAAACGCTTGTGAGCGTGAACCATGGCGAGACGAGCTG 1529
QY 501 ValThrAlaPro 504
DB 1530 GTGACAGCACCG 1541
RESULT 4
AR116881
LOCUS AR116881 1560 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 1 from patent US 6140042.
ACCESSION AR116881
VERSION AR116881.1 GI:14097787
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
1 (bases 1 to 1560)
AUTHORS Matsumoto, K. and Nishida, E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 6140042-A 1 31-OCT-2000;
FEATURES Location/Qualifiers
source 1..1560
/organism="unknown"
BASE COUNT 332 a 469 c 480 g 279 t

ORIGIN

Alignment Scores: 3.99e-151 Length: 1560
Pred. No.: 2580.00 Matches: 504
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 100.00% Gaps: 0
DB: 6

US-09-830-144-4 (1-504) x AR116881 (1-1560)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 30 ATGGCGCGCAGAGGAGGAGCTTGTCTGAGAGTGAGCAGCAGCCAAAGCTGGACAGATGAC 89
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTGGCTCAGCCCTCCAACCGCAGCTTCTGCTGAT 149
Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTyrLeuLysPheArgSerGluAsn 60
Db 150 GGCAAGGCACTGAGAGCCACCCCGCAGAGGACAGCTGGCTCAAGTTTCAGGAGTGAGAAC 209
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 270 GCCCAGCGGCTGTCCGAGAGCTCTCTGGCCAGCTGAATGCCGAGCAGCCGAGGCC 329
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 330 GATGTGGCGGTGTGTCTGTCAGAGGCTTCGATGTGGTGGAGGAGCTTCTTGGAGTCC 389
Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 390 ATTGACGACGCTTGGCTGAGAGGCAAGGCTCCAGTCGCAATTGCCAGAGGAGTCCCT 449
Qy 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 450 CAGCACCAGCTGCTCTCAGTATCAGAAATCCTTTGAGAGATCTCAAGACGTTAGAGAG 509
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 510 GAAATTTGGGAGGGGCGCATGGCGGTGTGGCGGCTCTTCTCAACAAACAGCTCTACGTC 569
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTCGGTACAAACCGTGACATTTTATGCAAAATCGACAGTGGATGGTTGCAGGTG 629
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 630 ACACAGCTGAACGTGGACCAACCCAGAGATGAGCTCTTCGCTCTTTCGCGAG 689
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyLysIleCysGlyGlnGluSerThr 240
Db 690 CTGGGCTTGGATGCTGAAAGATCAACAGGTGGGGATCATCTGTGGGCGAGGAGGACCC 749
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGGCGGATCGGGATTACAAGTTAAATATGGCTACACGAGCATTTGACCTTCTCAGCGCT 809
Qy 261 AlalysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGCGACAGCCGCTGGATGGG 869
Qy 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300
Db 870 GTGACGGGCTTCTTGGTGTGATCGGAGGGGTGTACAAAGGCCCTTAGAGGAGGCCCAT 929
Qy 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrThrGluPheAlaLysGln 320

Db 930 GGGCTGGCAGGCCAACCCAGGAGATTGCTGCGATGATTGACACTGAGTTTGCACAGCAG 989
Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
Db 990 ACCTCCCTGGCAGGAGTGCCAGGCGCTCTGTGACCGGTGAAGCGCATCCACAGCGAC 1049
Qy 341 ThrPheAlaSerGlyGlyGluArgAlaAraPheCysProArgHisGluAspMetThrLeu 360
Db 1050 ACCTTCGCCAGTGTGGGAGCGTGCAGGTTCTGCCCGCCGACGAGGACATGACCCCTG 1109
Qy 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
Db 1110 CTAGTGAGGAACCTTTGGCTACCCGCTGGGCGAAATGAGCCAGGCCACACCGAGCCGAGCC 1169
Qy 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCAGCTCGAGAGGAGAGTGATCCCTGTGTGTGCATACTCCAGCGCCAGAGCACC 1229
Qy 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1230 AGCAAGACCAGCGTGACCTCTCCCTTGTTCATGCCCTCCAGGGCCAGATGGTCAACGGG 1289
Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGTCTCACCTGGACGAAGCCACCCCGCCCTCACCAACCAAGCCCGACCC 1349
Qy 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460
Db 1350 TTAACCTTGCAGTCCACCAACACGACGACGAGCAGCAGCTCCAGCTCTGACGAGGCC 1409
Qy 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
Db 1410 CTCITTCGCTCCGCGCCGCCACTCGCTCCGCGCTGGCAGGAGCGTCTGTGTTGAGCCC 1469
Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
Db 1470 TATGTGGACTTGTGTGAGTTTATCCGCTCTGGAGCGTGGACCATGGCAGCAGAGCGTG 1529
Qy 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 5

E14752 Human mRNA for the TAB1 protein. 1560 bp DNA linear PAT 28-JUL-1999
LOCUS E14752
DEFINITION Human mRNA for the TAB1 protein.
ACCESSION E14752
VERSION E14752.1 GI:5709435
KEYWORDS JP 1998004976-A/1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1560)
AUTHORS Matsumoto, K. and Nishida, E.
TITLE TAB1 PROTEIN AND DNA CODING THE SAME
JOURNAL Patent: JP 1998004976-A 1 13-JAN-1998;
UENO NAOTO
COMMENT OS Homo sapiens (human)
PN JP 1998004976-A/1
PD 13-JAN-1998
PF 28-OCT-1996 JP 1996300856
PR 24-APR-1996 JP 96P 126282
PI MATSUMOTO KUNIHICO, NISHIDA EISUKE
PC C12N15/09, C07H21/04, C07K14/47, C07K19/00, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, G01N33/53, (C12N1/19, C12R1:865), (C12N1/21, C12R1:19),
PC (C12N5/10,
PC C12R1:91), (C12P21/02, C12R1:865), (C12P21/02, C12R1:91); CC
strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers

FH source 1..1560
 /organism='Homo sapiens'
 /tissue='kidney'
 /product='the TMB1 protein'
 /note='replace a'.
 FEATURES
 source 1..1560
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 BASE COUNT 332 a 469 c 480 g 279 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.99e-151 Length: 1560
 Score: 2580.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 6

US-09-830-144-4 (1-504) X E14752 (1-1560)

QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnGlnProSerThrAspAsp 20
 Db 30 ATGGCGGCGCAGAGGAGGAGCTTGCGAGAGTGAACAGCAGCAGCAGCATGAC 89
 QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
 Db 90 CTGCTCTCTGCGCACTCTCTGGGGTGGCTCAGCTCCAGCCGAGCTTCTCTGAT 149
 QY 41 GlyValGlyThrGlnSerHisProProGlnAspSerTyrLeuLeuPheArgSerGlnAsn 60
 Db 150 GCGAAGGCGCATGAGAGCACCAGCAGAGCAGCTGCTCAAGTTGAGAGTGAAGAC 209
 QY 61 AasnCyPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 269
 QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnAla 100
 Db 270 GCCGAGCGGCTCTCCGCAAGCTCTCTGGGCGCAGCTGAAAGCCGAGCAGCCGAGGCC 329
 QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGlnArgSerPheLeuGlnSer 120
 Db 330 GATGTGCGCGGTGTGTGCTGCGCAGGCTTCGATGTGTGAGAGAGAGGCTTCCTGAGTCC 389
 QY 121 IleAspAspAlaLeuAlaGlnLeuAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
 Db 390 ATTGACGAGCGCTTGCTGAGAGGCGAAGCCCTCCAGTGGCAATTGCCAGAGGAGATCCCT 449
 QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLeuThrLeuGlnArg 160
 Db 450 CAGCAGCAGCTGCTCTCTCAGATCAGAAAGATCCTTGAGAGCTCAAGACGTTAGAGAGG 509
 QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
 Db 510 GAAATTTTCGGAGGGGCGCATGCGCTTGCGGCTCTTCACCAACAAGCTCTACGTC 569
 QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
 Db 570 GCCAATGTGGTACAAACCGTGCCTTTATGCAAAATGACAGTGGATGGGTGGAGGTG 629
 QY 201 ThrGlnLeuAsnValAspHisThrThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
 Db 630 ACAAGCTGGAAGCTGGACCAACCAAGAGAGAGAGATGAGCTTCCGCTTTCCGAG 689
 QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCySerGlyGlnGlnSerThr 240
 Db 690 CTGGGCTTGATGTGGAAGAAGATCAAGCAGGTGGGATCATCTGTGGGCGAGAGACACC 749

QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db 750 CGGCGGATCCGGGATATCAAGATTAAATATGCTACACGAGCATTTGCTTCCAGCGCT 809
 QY 261 AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyAlaGlnProLeuAspGly 280
 Db 810 GCCAAGTCCAAACCAATCATGTGACAGACCCAGAAATCCATGGGGCACAGCCGCTGGATGGG 869
 QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaAlaHis 300
 Db 870 GTGACGGGCTTCTTGCTGCTGATGTGCGAGGGGTTGTACAGGCCCTTACAGGAGCCCAT 929
 QY 301 GlyProGlyGlnAlaAsnGlnGlnIleAlaAlaMetIleAspThrGlnPheAlaLysGln 320
 Db 930 GGGCTGGGAGGCCAACAGAGATTGCTGCATGATTGACACTGATGTTGCCAAGCAG 989
 QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
 Db 990 ACTTCCCTTGAGCGAGTGGCCAGGCCGTGTGTGACCGGGTGAAGCGCATCCACAGCGAC 1049
 QY 341 ThrPheAlaSerGlyGlyArgAlaArgPheCyPheProArgHisGlnAspMetThrLeu 360
 Db 1050 ACCTTGCCAGTGTGTGGAGCGCTGCCAGGTTCTGCCCCCGCAGAGAGATGACCTTG 1109
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGlnMetSerGlnProThrProSerProAla 380
 Db 1110 CTAGTGAGAACTTTGGCTACCCGCTGGGGCAATGAGCAGCCACACGAGCCAGCC 1169
 QY 381 ProAlaIleGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
 Db 1170 CCACTCCAGAGAGACGAGTGAACCTGTGTCTGTCTCCATCTCCAGCGCCAGACACC 1229
 QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
 Db 1230 AGCAAGACCAAGCGTGAACCTCTCTCTGTATGCTCTCCAGGCGCAGATGTGCACGGG 1289
 QY 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db 1290 GCTCACAGTCTTCCACCCCTGGAGCAAGCCACCCCACTCCCAACCAAGCCCGAGCC 1349
 QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSer 460
 Db 1350 TTAACTTCGACGTCCACCAACAGCAGCAGCAGAGAGAGAGCTTCCAGCTTGGAGGAGC 1409
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro 480
 Db 1410 CTCTTCGCTCCGCGCCGCGCCAGCTCGCTCCGCGCGGAGAGAGCGTGTGTAGGCC 1469
 QY 481 TyrValAspPheAlaGlnPheTyrArgLeuTyrPheSerValAspHisGlyGlnGlnSerVal 500
 Db 1470 TATGTGACTTGTCTGATTTACCGCTCTGGAGGTGAGCAATGCGAGCAGAGAGCGTG 1529
 QY 501 ValThrAlaPro 504
 Db 1530 GTGACAGCACCG 1541

RESULT 6
 LOCUS HSU49928 3096 bp mRNA linear PRI 06-APR-1998
 DEFINITION Homo sapiens TAK1 binding protein (TAK1) mRNA, complete cds.
 ACCESSION U49928
 VERSION U49928.1 GI:1401125
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 (bases 1 to 3096)
 Shibusawa, H., Yamaguchi, K., Shirakabe, K., Tonegawa, A., Gotoh, Y.,
 Ueno, N., Irie, K., Nishida, E. and Matsumoto, K.
 TAK1: an activator of the TAK1 MAPKK in TGF-beta signal
 transduction
 JOURNAL Science 272 (5265), 1179-1182 (1996)

LOCUS	AR058302	1560 bp	DNA	linear	PAT 29-SEP-1999
DEFINITION	Sequence 5 from patent US 5837819.				
ACCESSION	AR058302				
VERSION	AR058302.1 GI:5983879				
KEYWORDS	.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 1560)				
TITLE	Matsumoto, K. and Nishida, E.				
JOURNAL	TAB1 protein				
FEATURES	Parent: US 5837819-A 5 17-NOV-1998;				
source	Location/Qualifiers				
	1..1560				
	/organism="unknown"				
BASE COUNT	333 a	468 c	480 g	279 t	
ORIGIN					
Alignment Scores:					
Pred. No.:	8.14e-151	Length:	1560		
Score:	2575.00	Matches:	503		
Percent Similarity:	99.80%	Conservative:	0		
Best Local Similarity:	99.80%	Mismatches:	1		
Query Match:	99.81%	Indels:	0		
DB:	6	Gaps:	0		
US-09-830-144-4 (1-504) x AR058302 (1-1560)					
Oy	1	MetaIaAglARgArSserLeuEngInSerGIuGInGInProSerThrAspAsp	20		
Db	30	ATGGCGGCCCAAGAGAGAGCTTGTCTGCAGATGAGCACAACCAAGCTGACAGATGAC	89		
Oy	21	LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp	40		
Db	90	CTGCCTCTCTCGCACCTCTCTGGGGGTGGCTCAGCCTCCACCGCAGCTACTCTGCTGAT	149		
Oy	41	GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn	60		
Db	150	GGCAAGGGGACTGAGAGGCCACCGCCAGAGGACAGATGCTCAAGTTTCAGAGTGAAC	209		
Oy	61	AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrasnPheVal	80		
Db	210	AACCTGCTTCTGTATGGGCTTTCACACGGCTATGATGGCACAACGAGTACCAACTTCGTG	269		
Oy	81	AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla	100		
Db	270	GCCCGAGCGGCTGTCCGCAAGAGCTCTCTGGGCGCACCTCAATGCCAGACGCCGAGGCC	329		
Oy	101	AspValArgArgValLeuLeuEngInAlaPheAspValValGluArgSerPheLeuGluSer	120		
Db	330	GATGTGCGGCGGTGTGCTGTCTGCAAGGCTTCGATGTGTGAAGAGAGCTTCTTCGAGTCC	389		
Oy	121	IleAspAspAlaLeuValaGluLysAlaSerLeuEngInSerGlnLeuProGluGlyValPro	140		
Db	390	ATTGACGACCGCTTGCTGCTAGAAAGCAACCTTCACATTCGCAATGGCAAGGGAATGCCCT	449		
Oy	141	GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg	160		
Db	450	CAGCACACCACTGCTCTCCTAGTATCAGAAAGATCTCTTGAAGAAGCTCAAGACGTTAAGAGAG	509		
Oy	161	GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal	180		
Db	510	GAAATTTCCGGAGGGGCCATAGCCGTTGGCGGCTCTTCTCAACAACAAGCTCTACGTC	569		
Oy	181	AlaAsnValGlyThrasnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal	200		
Db	570	GGCAATGTTCGTAACAACCGTGCACCTTTATGCAATTCACAGTGGATGGGTTCAGAGTG	629		
Oy	201	ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln	220		

Db	630	ACACAGCTGAACGTGGACACACACACAGAGAACGAGGTGAGCTCTTCCGCTTCCGAG	689
Qy	221	LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGlnSerThr	240
Db	690	CTGGGCTTGGATGCTGGAAAGATCAACACAGGTGGGATCATCTGTGGGACAGAACGCC	749
Qy	241	ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerLeu	260
Db	750	CGGGGAGTCGGGGATTCACAGGTTAAATATGGCTACACGGACATTGACTTCTCAGCCCT	809
Qy	261	AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyValGlnProLeuAspGly	280
Db	810	GCCAAAGTCCAAACCAATCATTCGAGAGCCAAATTCATGGGGCACAGCCGCTGGATGG	869
Qy	281	ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis	300
Db	870	GTCAGCGGCTTCTTGGCTGATGTCCGAGGGGTTGTAACAAGGCCCTAAGAGCGAGCCAT	929
Qy	301	GlyProGlyGlnAlaAsnGlnGlnIleAlaIleMetIleAspThrGlnPheAlaLysGln	320
Db	930	GGGGCTGGCGAGGCCAACCAAGAAATTGCTGCATGATTGACACTGAGTTTGGCCAAAGAG	989
Qy	321	ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp	340
Db	990	ACCTTCCTGGACGCGAGTGGCCCAAGCCGCTCGTGGACCGGGTGAAGCGCATCCACAGCAC	1049
Qy	341	ThrPheAlaSerGlyGlyLysArgAlaArgPheCysProArgHisGlnAspMetThrLeu	360
Db	1050	ACCTTCGCGCAATGTGTGGGAGCGTGCACAGTTCTGCCCGCCGCGACGAGGACATGACCCTG	1109
Qy	361	LeuValArgAsnPheGlyTyrProLeuGlyGlnMetSerGlnProThrProSerProAla	380
Db	1110	CTAGAGAGAACTTGGCTGACCCGCTGGCGAAATGACGCCACCCACACCGAGCCACGCC	1169
Qy	381	ProAlaIleGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr	400
Db	1170	CCAGCTGCAGAGACGACGAGTTCACCTGTGTGTGTCATCTCCAGGCCACAGACACC	1229
Qy	401	SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly	420
Db	1230	AGCAAGACCAACGTCGACCTCTCTTGTATGCCCTCCAGGGCCAAATGTCACACGGG	1289
Qy	421	AlaHisSerAlaSerThrLeuAspGlnAlaArgProThrLeuThrAsnGlnSerProThr	440
Db	1290	GCTCACTGTGCTTCACCCCTGAGCAAGACCAACCCCACTTACCAACCAAGCCCGACCC	1349
Qy	441	LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGly	460
Db	1350	TTAACCTTCGCACTCACCACCAACGACACGACGACGACGACTCCAGCTTCGACGGAAGC	1409
Qy	461	LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro	480
Db	1410	CTCTTCGCTCCGCGCCGCGCCCACTCGCTCCGCTCGCGCGAGAGAGCGTGGTTAGAGCC	1469
Qy	481	TyrValAspPheAlaGlnPheTyrArgLeuTyrSerValAspHisGlyGlnGlnSerVal	500
Db	1470	TATGGAGCTTGTGCAATTTTACCGCTTCGAGCGGTGAGCAATGGGAGACAGAGCTGG	1529
Qy	501	ValThrAlaPro 504	
Db	1530	GTCAGACGACCG 1541	
RESULT 8			
LOCUS	AR088276	1560 bp	DNA linear PAT 07-SEP-2000
DEFINITION	Sequence 5 from parent US 5989862.		
ACCESSION	AR088276		
VERSION	AR088276.1	GI:10015039	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1560)		

AUTHORS Matsumoto, K. and Nishida, E.
TITLE Tab1 protein and DNA coding therefor
JOURNAL Patent: US 5989862-A 5 23-NOV-1999;
FEATURES Location/Qualifiers
source 1..1560

BASE COUNT	333 a	468 c	480 g	279 t
ORIGIN				

Alignment Scores:	
Pred. No.:	8, 14e-151
Score:	2575.00
Percent Similarity:	99.40%
Best Local Similarity:	99.80%
Query Match:	99.01%
DB:	6
Length:	1560
Matches:	503
Conservative:	0
Mismatches:	1
Indels:	0
Gaps:	0

US-09-830-144-4 (1-504) x AR088276 (1-1560)

Qy	1	MetAlaAlaGlnArgArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp	20
Db	30	ATGCGCGCAGAGGAGGAGCTGCTCAGAGTGAGCAGCAGCAAGCTGGACAGATGAC	89
Qy	21	LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp	40
Db	90	CTGCCTCTCTGCCACCTCTCTGGGTGGCTCAGCTCCAACCGCAGCTACTCTGCTGAT	149
Qy	41	GlyIysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn	60
Db	150	GGCAAGGGCACTGAGAGCCACCCGCCACAGACAGATGGCTCAAGTTCAGAGTGAGAAC	209
Qy	61	AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal	80
Db	210	AACTGCTTCCTGTATGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCACTTCGTG	269
Qy	81	AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla	100
Db	270	GCCACAGCGGTGTCGCGAGAGCTCTCGTCGGCCAGCTGAATGCCAGCACGCCGAGGCC	329
Qy	101	AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer	120
Db	330	GATGTGCGGGCTGTGCTGTGACGGCTTCGATGTGTGGAGGAGCTTCCTGGAGTCC	389
Qy	121	IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro	140
Db	390	ATTGACGACGCCCTTGGCTGAGAAGGCAAGCTCCAGTCGCAATTGCCAGAGGGAGTCCCT	449
Qy	141	GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg	160
Db	450	CAGCACGAGCTGCCCTCTCAGTATCAGAAGATCCTTGAGAGACTCAAGACCTTAGAGAGG	509
Qy	161	GluIleSerGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal	180
Db	510	GAATTTTCGGAGGGGCCATGGCCGTTGTGGCGCTCTCTCAACCAACAAAGCTCTACGTC	569
Qy	181	AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal	200
Db	570	GCCAAATGCGGTACAAACCGTGACATTTTATGAAATCGACAGTGATGGGTTCAGAGTG	629
Qy	201	ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln	220
Db	630	ACACAGCTGAACGTGGACCACACACAGAGAACCGAGATCAGCTCTTCGTCCTTCGCGAG	689
Qy	221	LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr	240
Db	690	CTGGGCTTGTATGCTGAAAGATCAACAGGTGGGATCATCTGTGGGAGGAGAGACACC	749
Qy	241	ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla	260
Db	750	CGCGCGATCGGGATTAACAAGGTTAAATATGGCTACCGGACATTGACCTTTCACGGCT	809
Qy	261	AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly	280

Db	810	GCCAGTCCAAACCAATCATCGACAGCCAGAAATTCATGGGGCACAGCCGCTGGATGGG	869
Qy	281	ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHis	300
Db	870	GTGAGGACTTCTTGGTGCTGATGTCGAGGGGTGTACAGGCCCTAGAGGCACCCCAT	929
Qy	301	GlyProGlyGlnAlaAsnGlnGluLeuAlaMetIleAspThrGluPheAlaLysGln	320
Db	930	GGGCTGGGACAGGCCAACACAGAGATTCTCGATGATTGACACTGAGTTTGCCCAAGCAG	989
Qy	321	ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp	340
Db	990	ACCTCCCTGGACAGAGTGGCCAGGCCCTGTGGACCGGGTGAACGGCATCCACAGCGAC	1049
Qy	341	ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu	360
Db	1050	ACCTTCGCCAGTGGTGGGAGCGTGCCAGGTTCTGCCCCCGGCACGAGGACATGACCCCTG	1109
Qy	361	LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla	380
Db	1110	CTAGTGAGGAATTTGGCTACCCGCTGGCGCAATAGCCAGCCACACCGAGCCAGCC	1169
Qy	391	ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr	400
Db	1170	CCAGCTGCAGAGGACGAGTGTACCTGTGTCTGTGCCATATCTCAGCGCCACAGACACC	1229
Qy	401	SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly	420
Db	1230	AGCAAGACCAGGTGACCCCTCTCCCTTGTCTATGCCCTCCAGGGCCAGATGTTCAACGGG	1289
Qy	421	AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr	440
Db	1290	GCTCACAGTGTCTCCACCCTCGAGAACCCACCCTCACCAACCAAGCCCGACC	1349
Qy	441	LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly	460
Db	1350	TTAACCTTCGAGTCCACCAACGACACGACGACGAGCAGCAGCTCTCAGGTCCTCAGCGAGC	1409
Qy	461	LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro	480
Db	1410	CTTTTCCTCCCGCCGCCACTCGCTCCGCTCGCGAGGACGGTCTGTGAGCCCC	1469
Qy	481	TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGlnSerVal	500
Db	1470	TATGTGACTTTGTGAGTTTACCGCTCTGGAGCGTGGACATGGCGCAGCAGCGCTG	1529
Qy	501	ValThrAlaPro 504	
Db	1530	GTGACAGCACCG 1541	
RESULT	9		
ARL16884			
LOCUS	ARL16884	1560 bp	DNA linear PAT 16-MAY-2001
DEFINITION	Sequence 5 from patent US 6140042.		
ACCESSION	ARL16884		
VERSION	ARL16884.1	GI:14097790	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1560)		
AUTHORS	Matsumoto, K. and Nishida, E.		
TITLE	TAB1 protein and DNA coding therefore		
JOURNAL	Patent: US 6140042-A 5 31-OCT-2000;		
FEATURES	Location/Qualifiers		
source	1..1560		
	/organism="unknown"		
BASE COUNT	333 a 468 c 480 g 279 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	8.14e-151	Length:	1560
Score:	2575.00	Matches:	503

Percent Similarity: 99.80% Conservative: 0
 Best Local Similarity: 99.80% Mismatches: 1
 Query Match: 99.81% Indels: 0
 DB: 6 Gaps: 0

US-09-830-144-4 (1-504) X AR116884 (1-1560)

QY 1 MetAlaAlaGlnArgArgSerLeuLeuGlnSerGlnGlnProSerTrpThrAspAsp 20
 Db ATGGCGGCGGAGAGGAGAGAGCTTGTCCAGAGTGCACAGCAGCAGCAGCTGGACAGTGC 89
 QY 21 LeuProLeuCySHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
 Db CTGCTCTCTCTGCCACTCTCTTGGGGTTGGCTCAGCTCCACCGGAGCTACTCTGTGAT 149
 QY 41 GlyLysGlyThrGlnLeuSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
 Db GGCAAGGGGCACTGAGAGCCACCGCCAGAGGACAGATGGCTCAAGTTCCAGAGTGAGAAC 209
 QY 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db AACTGCTTCTCTATGAGGGGCTTCAACGGCTATGATGGCAACCGAGTGCACAACTTCGG 269
 QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlnGlyGlnLeuAsnAlaGlnHisAlaGlnAla 100
 Db GCGCAGGCGGCTGTCCGAGAGCTCTGCTGGCCAGCTCAATCCAGACCGCCAGAGGCC 329
 QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
 Db GATGTGGGGGTGGTGGCTGAGAGCCCTTGATGTGTGGAGAGGAGCTTCCCTGGAGTCC 389
 QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGluValPro 140
 Db ATTGACGACGCCCTTGGTGAGAGGCAAGCTTCAGTCCGAAATTCGACAGGAGAGCCCT 449
 QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArg 160
 Db CAGACCAAGCTGCTCTCTCAGTATCCAGAGATCTCTTGAGAGATCTCAAGACGTTAGAGAG 509
 QY 161 GlnIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
 Db GAATATTCGGGAGGGGCGCATGGCGGTGGCGCTCTTCTCAACAACAAGCTCTACGTC 569
 QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
 Db GCCAATCTCCGTACAAACCGTCGACTTTATGCAAAATCGACAGTGGATGGTTGCGAGTG 629
 QY 201 ThrGlnLeuAsnValAspHisThrThrGlnAsnGluAspGluLeuPheArgLeuSerGln 220
 Db ACACAGCTGAACCTGAGCCACACACACAGAGATGAGCTTTCCTCTTCCGTCAG 689
 QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCySglnGlnSerThr 240
 Db CTGGGCTTGATCTGGAAGAAGATCAAGCAGGTGGGATCTGTGGGCGAGGAGACACC 749
 QY 241 ArgArgGlyLeuLysPyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db CGCGCGGATCCGGGATTACAAAGGTTAATATGCTACACGACATTACCTTCCACGCT 809
 QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
 Db GCCAAGTCCAAACCAATCATTCGACAGCCAGAAATCATGGGCGACAGCCGCTGGATGG 869
 QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaAlaHis 300
 Db GTAGCGGGCTTCTTGGTGTGATGTGGAGGGGTGGTGAAGGCCCTTAAAGGCGAGCCAT 929
 QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
 Db GGGGCTGGGAGGCGCAACGAGATTTGTCGCACTGATGACTAGTTTGGCAAGCAG 989
 QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340

Db 990 ACTCCCTGAGCGCAGAGTGGGCCAGAGCCGTGTG9AACCGGGTGAAGCGCATCCACAGGAC 1049
 QY 341 ThrPheAlaSerGlyGlyGluValGalaArgPheCyProArgHisGlyLysAspMetThrLeu 360
 Db ACTTTCGCGCAGTGTGGGGAGCGGTGCGAGGTTCTGCCCCCGGAGCAGGAGCAATGACCTTG 1109
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db CTAGTAGAGAACTTTGGCTTACCGCTGGCGAATGAGCCAGCCACCGAGCCAGGCC 1169
 QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
 Db CCAGCTGGAGAGAGCAGAGTGAACCTGTGTGTGTCATCTCCAGCTCCAGCGCCAGAGACC 1229
 QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
 Db AGCAAGACCAAGCGTACCCCTCTCCCTTGTGATGCTCCCGCAGGCGCAGATGGTCAACGGG 1289
 QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db GCTCACAGTGTCTTCCACCTGAGAGAACCAACCCACCTTCACCAACCAAGCCGAGACC 1349
 QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460
 Db TTAAACCTGCGAGTCCACCAACAGCAGCAGCAGAGCAGAGCTTCAGCTTGAAGAGGC 1409
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyValGluPro 480
 Db CTCTTCCGCTCCCGCGCCGCGCTCCCTCCGCTGGCAGAGCGCTCTGTGTGAGCCCC 1469
 QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGlnGlnSerVal 500
 Db TATGTGACCTTGTCTGATGATTTTACCCTCTTGAGCTGAGCATGGCAGACAGAGCTG 1529
 Db 1470 TATGTGACCTTGTCTGATGATTTTACCCTCTTGAGCTGAGCATGGCAGACAGAGCTG 1529
 QY 501 ValThrAlaPro 504
 Db 1530 GTGACAGCACCG 1541

RESULT 10

BC027054
 LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC027054 2944 bp mRNA linear ROD 07-AUG-2002
 Mus musculus, Similar to mitogen-activated protein kinase kinase
 kinase 7 interacting protein 1, clone IMAGE:504323, mRNA, partial
 cde.
 BC027054
 BC027054.1 GI:20072133
 house mouse.
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2944)
 Strausberg, R.
 Direct Submission
 Submitted (04-APR-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 45 Row: k Column: 11
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, GenesScan gene prediction.

FEATURES

source

Location/Qualifiers

1..2944

/organism="Mus musculus"

/db_xref="taxon:10090"

/map="FVB/N"

/clone="IMAGE:504323"

/issue_type="Kidney, normal. 5 month old male mouse."

/clone_lib="NCI CGAP_Kid14"

/lab_host="DH10B"

/notes="Vector: pCMV-SPORT6"

<1..1503

/codon_start=1

/product="Similar to mitogen-activated protein kinase

kinase kinase 7 interacting protein 1"

/protein_id="AAH27054.1"

/db_xref="GI:20072134"

/translation="AQRSLLOSQQPSTWTDPLCHLSGVGSASNRYSADGKTES
HPEDNWLKRSNNCFYGVNGYDGRVTFVQRLSAELLGLQLTEADVR
VLLQPDVVERFLESIDDAEAKSLQSLQEPGVQQLPPQYQILERLKAUREI
SGAMAVAVLLNSKLYVANVTNRALLCKSTVDGLVTQLNMDHTNEDELFRSLQ
LGLDAGIKQMGYICQESTRRIGYKVKYGYTDIDLLSAAKSPKPIIAEPETHGAQPL
DGVTFGLVMSGLYKALEAAHPGQANQEIAMIDTFEAKTQSDVAQVAVRVRK
IHSDTFASGGERAKFCPRHEDMTLLVRNFGYPLGBMSOPTPTPAPGRVVPVPSYS
AOSTKSTVTLVMPSSQOMVNGSHSASTLDEATPTLNTQSPITLIGSTNHTHQSSS
SSSDGLFRSRFAHSLPPGEDGRVFPYDFAEFRLMSVDHGEQSVMTAP"

BASE COUNT 676 a 835 c 863 g 570 t

ORIGIN

Alignment Scores:

Pred. No.: 1,44e-146 Length: 2944
Score: 2512.00 Matches: 488
Percent Similarity: 99.00% Conservative: 9
Best Local Similarity: 97.21% Mismatches: 3
Query Match: 97.36% Indels: 2
DB: 10 Gaps: 1

US-09-830-144-4 (1-504) x BC027054 (1-2944)

QY 3 AlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspLeuPro 22
Db 1 GCGCAGAGGAGGACCTGCTCCAGAGTGACGACGACCAAGCTGGACAGTGAACCTGCCA 60
QY 23 LeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLys 42
Db 61 CTCTGTCACTCTCTGGAGTTGGTTTCAGCGTCCAAACCGCAGCTACTCTGCTGATGGCAAG 120
QY 43 GlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsnAsnCys 62
Db 121 GGCACTGAGAGGACCACTCCCGAGGACAACTGGCTTAAGTTCAGAAAGTGAATAACTGC 180
QY 63 PheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGln 82
Db 181 TTCTGTACGGGGTCTTCAATGGCTATGATGGCAACCGCGTGACCAACTTTGTGGCGCAG 240
QY 83 ArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAlaAspVal 102
Db 241 AGGCTTTCTGAGAGCTCTCTGCTGGGCGAGCTCAACTGAGCACACTGAGGCTGATGTG 300
QY 103 ArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSerIleAsp 122
Db 301 CGACGGGTCTGTCTGACGGCTTCGATGTGTGGAGAGGAGCTTCCTGGAGTCTATCGAT 360
QY 123 AspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHis 142
Db 361 GATGCCCTTAGCTGAGAAGCAAGCCTCCAGTCCAGCTCCGAGGGGTGTGCCCCAGCAC 420
QY 143 GlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIle 162

Db 421 CAGCTGCCACCTCAGTACCAGAGATCCTTGAGAGATCAAGGCACCTGGAGAGGAGATT 480
QY SerGlyGlyAlaMetAlaValValAlaValLeuLeuAsnAsnLysLeuTyrValAlaAsn 182
Db 481 TCGGAGGAGGACCATGGCTGTGTGGCAGTCTTCTCAACAGCAAGCTCTACGTGGCAAT 540
QY ValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGln 202
Db 541 GTCGGTACGAACAGGGCCCTTCTGTGCAAAATCTACAGTGGATGGTTTACAGGTTACACAG 600
QY LeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGly 222
Db 601 CTAACATATGGACACACCGAGACGAGCAGAGTCTTTTTCGGCTTTCGCAACTGGT 660
QY LeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGluSerThrArgArg 242
Db 661 TTAGACGAGGAGAGATCAAGCAGATGGCGTCTATCTGTGACAGAGAGACACCGCGG 720
QY IleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLys 262
Db 721 ATTGGGATTAAGGTCAAAATATGGCTACACCGACATTCACCTGTCTCAGCGCTGCCAAG 780
QY SerLysProIleIleAlaGluProGluIleHISGlyAlaGlnProLeuAspGlyValThr 282
Db 781 TCCAAACCATCATCGCAGACCGGAAATTCATGGCCACAGCCTCTGGATGGCGTGACG 840
QY GlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyPro 302
Db 841 GGCTTCCTGGT 900
QY GlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSer 322
Db 901 GGGCAGGCGCAACAGGAGATTGCCGCGATGATTGACACCGAGTTTGCACAGCAGACCTCC 960
QY 323 LeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAspThrPhe 342
Db 961 CTGGATGCGGTGTGCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 343 AlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeuLeuVal 362
Db 1021 GCCAGTGTGGGGAGCGTGCCTCAAGTTCTGCCCCAGGATGAGACATGACCTCTCTGTGT 1080
QY 363 ArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAlaProAla 382
Db 1081 AGGAATTTGGCTATCCATTGGGTGAAATGAGCCAGCCTACACCGACCCCTGCCCA --- 1137
QY 383 AlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThrSerLys 402
Db 1138 ---GGGGGCCGTGTGTACCTCTTCTGTGCCCTACTCAAGTGCCCGAGACACCGCAAG 1194
QY 403 ThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGlyAlaHis 422
Db 1195 ACTAGTGTACTCTGTCTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCT 1294
QY 423 SerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThr 442
Db 1255 AGTGCTCCACTCTGGACGAAGCCACTCCACACTCACTAACACGAGAGCCCACTCTGACC 1314
QY 443 LeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSerSerSer 462
Db 1315 CTGCAAGTCCCAAC 1374
QY 463 ArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluProTyrVal 482
Db 1375 CGCTCCAGACCGGCTACTCTACTCTCCACCGGAGAGGATGGCCGTGTGGAGCCCTATGTG 1434
QY 483 AspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerValValThr 502
Db 1435 GACTTTGTGTGAGTCTACCGACTCTGGAGCGTGGNACCAGCGGAGAGCGGTGATGACG 1494
QY 503 AlaPro 504
|||||

Db 1495 GCACCT 1500

RESULT 11
LOCUS XLU92031 1926 bp mRNA linear VRT 18-APR-1998
DEFINITION Xenopus laevis TAK1 binding protein TAB1 mRNA, complete cds.
ACCESSION U92031
VERSION U92031.1 GI:3057037
KEYWORDS
SOURCE Xenopus laevis.
ORGANISM Xenopus laevis.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 1926)
Shibuya,H., Iwata,H., Masuyama,N., Gotoh,Y., Yamaguchi,K., Irie,K.,
Matsumoto,K., Nishida,E. and Ueno,N.
Role of TAK1 and TAB1 in BMP signaling in early Xenopus development
EMBO J. 17 (4), 1019-1028 (1998)
JOURNAL
MEDLINE 98130593
PubMed 9463380
REFERENCE
2 (bases 1 to 1926)
Shibuya,H.
Direct Submission
Submitted (05-MAR-1997) Faculty of Sciences, Hokkaido University,
Nishi 6-chome, Kita 12, Kita-ku, Sapporo, Hokkaido 060, Japan
FEATURES
source
1..1926
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
38..1534
/note="TAK1 binding protein"
/codon_start=1
/product="TAB1"
/protein_id="AAC14009.1"
/db_xref="GI:3057038"
translation="MAAPRRNLHSQSWTDLPICNLISVGASASNOTYNSRGLKDEH
PYEDWIKFRGDNNTLYGVNRYEGTATSFVGRORLAELLGLDPTDVAHVHY
LQAFDVVERFLESIDDLAEKASLSQLEPGLHQLPQYQKIVDRNLTKELY
GGAMVIVLVNSKLVANVNGTRALILCKSTVDGIVQVLADHTENEDILLSOL
GLDPTIKIQOVYIGGQSTRIGDYKVNPNDELELSTASKSPITAEPEHGCQPLD
GVDTFLVMSSEGLYKALSAHPGOANOELAMLTAEPAKQVSDVEVAALVEEVKSI
HNDTPASGERAKYCSKHEMDTLVRNGYPLQETSPPTLPTGGRIKPYPSVSSA
QNTSKTSVTLISLWPSQGPVNVGNSSSTLDGTTSTLOSPEATLOSTRITRQSSSSS
DGGFRSRPLPLQPEDGKVEPVDFTFYRLWMAERNDPTLTAQ"

BASE COUNT 556 a 462 c 456 g 451 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	2,14e-110	Length:	1926
Score:	1924.00	Matches:	378
Percent Similarity:	85.71%	Conservative:	54
Best Local Similarity:	75.00%	Mismatches:	64
Query Match:	74.57%	Indels:	8
DB:	5	Gaps:	4

US-09-830-144-4 (1-504) x XLU92031 (1-1926)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnGlnInProSerTrpThrAspAsp 20
Db 38 ATGGCTGGCCCTAGAGGAATCTTCTGCATAGTCAG-----AGCTGCAACATGAT 88

Qy 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTrpSerAlaAsp 40
Db 89 CTCCTCTCTGCAACTTCTCTGGGGTTGGATCTGCTTCCATCAGACTTACCACTCAGAG 148

Qy 41 GlyValSerGlyThrGluSerHisProGluAspSerTrpLeuLeuPheArgSerGluAsn 60
Db 149 GGGCTTTGGGAAAGCAGACCATTAAGAATACTGATTAAGTTCAAGGGGCGACCAT 208

Qy 61 AsnCySPheLeuTyGlyValPheAsnGlyTyAspGlyAsnArgValThrAsnPheVal 80
Db 209 AATATTATCTTTATGAGAGTGTTTAAAGNGTATGAGGGAACCAAGCCACGCTTTGTG 268

Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 269 GGTCAAGCGCTTGACGACGAGAACTTCTTGCGGCGAGTTGGACCTTATGCACAGATGCA 328

Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSer 120
Db 329 GAGGTCCAAAGATATTGCTACAGGCACTTGATGTTGTTGAAGAAAGCTTTCGGAGTCC 388

Qy 121 IleAspAspAlaLeuAlaGluValAspSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 389 ATCGATGATTGCTTAGCAGAAAAGCCAGCTTCTGTCCTCCACTTCCAGAGGCTTACTCTG 448

Qy 141 GlnHisGlnLeuProProGlnTyTrpGlnValIleLeuGlnArgLeuTySerThrLeuGluArg 160
Db 449 CACCAACACCTTCCAAACCCAGTCCAAAGAAATGTTGTGACCGGCTTAAACATATCTTGAGAG 508

Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnValLeuTyTrpVal 180
Db 509 GAATATATGGGGAGGCGCATGTCATTTGTGCTTATTTGTGAACAGCAACTTATGTT 568

Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLeuSerThrValAspGlyLeuGlnVal 200
Db 569 GCAATGTCCGAAACAACACAGACACTTTGTGTAAATCACTTAGATGGCTTCAAGTC 628

Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 629 ACTCAGCTCAATGCTGACACACACTACCGAAATGAAGATGAGATATGCGTGTCTCAA 688

Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuGlyGlnGlnSerThr 240
Db 689 TTGGGCTGGACACTACAAAGATTAAACAGGTGGGGTTTGTGAGGCGCAACACACT 748

Qy 241 ArgArgIleGlyAspTrpLysValLysTyTrpGlyTyThrAspIleAspLeuLeuSerAla 260
Db 749 CGCAGATTGGAGACTCAAGAGCTCAAGGTCAAGTCAACTCAATGACATTGAGCTGCAACACA 808

Qy 261 AlaLysSerLysProIleIleAlaGluProGlnIleHisGlyValaGlnProLeuAspGly 280
Db 809 GCCAAGTCCAAAGCTATTTACTGCAAGCCAGAGATCCATGCTGCCAGCATTAAGCGGT 868

Qy 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyTrpValaLeuGlnAlaHis 300
Db 869 GTCCAGAGATTTTGGTCTTAATGTCTGAAGACCTTTCAGAGGCTCTCCAGTACGCCAT 928

Qy 301 GlyProGlyAlaAsnGlnIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
Db 929 GGGCTGGAACGGCCAAACGAAATAGCGCATGATTCGCCAGAGTTGCGCAAGCAG 988

Qy 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
Db 989 GTTCTCTCTGATGAGTGGCAGCAGGCTTGTGTGAGAGGGTTAAACGGATTTCCACATGAT 1048

Qy 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCySPheArgHisGluAspMetThrLeu 360
Db 1049 ACGTTTCAAGTGTGGCGAAGCGGCAAAATATTGAGTAAAGCAACAGACAGACGCTA 1108

Qy 361 LeuValArgAsnPheGlyTyTrpLeuGlyGluMetSerGlnInProThrProSerProAla 380
Db 1109 CTGTGCGGAACCTGGGCTTACCGGCTTCAGAGATACAGCCCCCACCACACTCAGC----- 1162

Qy 381 ProAlaAlaGlyArgValTyProValSerValProTySerSerAlaGlnSerThr 400
Db 1163 CCCACTCAAGGTGAGCCTTATATCCAGTGTCTGTGCCATATTCAGATGCTCCAGATACC 1222

Qy 401 SerTyThrSerValThrLeuSerLeuValMetProSerGlnGlyGluMetValAsnGly 420
Db 1223 AGTAAACAAAGTGCACACTGCTACGTGATGCCATCAAGGTTCCAAATGGGATGGA 1282

Qy 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
Db 1283 ACAAAACGACGCTCAACACTGGATGGAAACAACGTCACACTG-----CAATCTCCAACT 1336


```

QY 441 LeuThrluGlnSerThrAenThrHisThrGlnSerSerSerSerSerAepGlyGly 460
      |||||
Db 1337 GCTACTCTCCAGTCCACCACTCACACTCAAAGCAGCAGTCCAGTCTCTGATGAGGC 1396
      |||||
QY 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspClyArgValGluPro 480
      |||||
Db 1397 CTGTTCCGATCTCGCCATTCGCTCTTTGCACCCAGATGAGGATGCCGCGTGGAACCC 1456
      |||||
QY 481 TyValAspPheAlaGluPheTyArgLeuTyrSerValAspHisGlyGluGln--Ser 499
      |||||
Db 1457 TACGTTGACTTCACAGACTTTTACCGCTTTGGAATGCAGACATATGATCCGGGACT 1516
      |||||
QY 500 ValValThrAla 503
      |||||
Db 1517 CTGCTCACTGCA 1528

RESULT 12
HS407F17 69660 bp DNA linear PRI 12-DEC-1999
LOCUS Human DNA sequence from clone RP3-407F17 on chromosome 22 Contains
DEFINITION the gene for TAB1 (TAK1 binding protein 1), ESTs, STSs, GSSs and
SOURCE two putative CpG islands, complete sequence.
ACCESSION Z83845
VERSION Z83845.14 GI:5441636
KEYWORDS HTG; CpG island; TAB1; TAK1.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 69660)
Corby,N.
Direct Submission
Submitted (08-DEC-1999) Sanger Centre, Hinxton, Cambridgeshire,
CE10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
This sequence was generated from part of bacterial clone contigs of
human chromosome 22, constructed by the Sanger Centre Chromosome 22
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr22
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/Celegans/wormpep
from the library RPCI-3 constructed at the Roswell Park Cancer
Institute by the group of Pieter de Jong. For further details see
http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
IMPORTANT: This sequence is not the entire insert of clone
RP3-407F17. It may be shorter because we only sequence overlapping
sections once, or longer because we arrange for a small overlap
between neighbouring submissions.
The true left end of clone RP3-407F17 is at 1 in this sequence. The
true left end of clone RP5-1104E15 is at 69561 in this sequence.
The true right end of clone RP3-333H23 is at 17123 in this
sequence. The start of this sequence overlaps with sequence
AL022326..

Location/Qualifiers
1..69660
/organism="Homo sapiens"

FEATURES
Source

```

```

/db_xref="taxon:9606"
/chromosome="22"
/clone="RP3-407F17"
/clone_lib="RPCI-3"
30..216
/notes="AluJb repeat: matches 111..296 of consensus"
31..412
/notes="match: STS: Em:HS407F17S; match: STS: Em:AL021668"
346..437
/notes="MER45B repeat: matches 1..74 of consensus"
438..753
/notes="AluSx repeat: matches 1..308 of consensus"
754..790
/notes="MER45B repeat: matches 74..114 of consensus"
879..1189
/notes="AluSq repeat: matches 1..313 of consensus"
1198..1507
/notes="AluSg repeat: matches 1..298 of consensus"
1855..2147
/notes="AluSx repeat: matches 1..293 of consensus"
2367..2845
/notes="match: GSS: Em:AQ877061"
2485..2857
/notes="AluSg/x repeat: matches 135..308 of consensus"
3146..3341
/notes="MER45B repeat: matches 862..1039 of consensus"
3504..3817
/notes="AluY repeat: matches 1..311 of consensus"
3820..3836
/notes="MIR repeat: matches 193..207 of consensus"
3837..4187
/notes="AluSx repeat: matches 1..312 of consensus"
4188..4356
/notes="MIR repeat: matches 8..193 of consensus"
4613..4797
/notes="AluSg/x repeat: matches 126..306 of consensus"
4864..5019
/notes="L2 repeat: matches 1599..1763 of consensus"
5026..5206
/notes="L2 repeat: matches 2241..2420 of consensus"
5233..5594
/notes="AluYb8 repeat: matches 1..308 of consensus"
5537..5890
/notes="AluJb repeat: matches 12..303 of consensus"
5891..6057
/notes="AluSp repeat: matches 126..300 of consensus"
6068..6236
/notes="L2 repeat: matches 2544..2750 of consensus"
6539..7149
/notes="L2 repeat: matches 1815..2541 of consensus"
7371..7432
/notes="MIR repeat: matches 158..219 of consensus"
7439..7516
/notes="AluJb repeat: matches 1..60 of consensus"
7517..7691
/notes="AluSc repeat: matches 132..306 of consensus"
7692..7986
/notes="AluSg repeat: matches 1..294 of consensus"
7987..8213
/notes="AluJb repeat: matches 53..281 of consensus"
8297..8506
/notes="AluSp repeat: matches 2..212 of consensus"
8691..8994
/notes="AluSx repeat: matches 1..301 of consensus"
9055..9367
/notes="AluSg repeat: matches 1..310 of consensus"
9532..9655
/notes="AluSg/x repeat: matches 69..192 of consensus"
9656..9804
/notes="AluJb repeat: matches 149..297 of consensus"
9873..10415
/notes="CpG island"
/evidence=not_experimental

```

gene
10049. .42131
/gene="d407f17.1"
Join(10049. .10084,25255. .25391,25749. .25902,27040. .27126,
27960. .28098,28981. .29094,29768. .29879,32076. .32220,
36952. .37174,38270. .38432,40264. .42131)
/gene="d407f17.1"
/product="d407f17.1 (TAB1 (TAKI binding protein 1))"
/note="match: CDNAS: Em:U92031 Em:U49928; match: ESTs:
Em:H17295 Em:AA726609 Em:AA019355 Em:AA168567 Em:AE54435
Em:AA789149 Em:AI014848 Em:AA430502 Em:AA745822
Em:AA411895 Em:AA411894 Em:AA326751 Em:AA430621 Em:HS9337
Em:AA430462 Em:AA806521 Em:R40486 Em:AA019356 Em:AA157525
Em:AA028980 Em:T34667 Em:R83147 Em:AI269211 Em:AA86517
Em:AA147685 Em:AA458466 Em:243346 Em:AI656463 Em:W95352
Em:AA828095 Em:R60090 Em:AA280338 Em:AA679208 Em:AA442978
Em:FO8391 Em:AI085576 Em:W95394 Em:R13055 Em:H48607
Em:R60028 Em:FO2762"
/evidence=not experimental
Join(10052. .10084,25255. .25391,25749. .25902,27040. .27126,
27960. .28098,28981. .29094,29768. .29879,32076. .32220,
36952. .37174,38270. .38432,40264. .40471)
/gene="d407f17.1"
/note="match: proteins: Tr:Q15750 Tr:O73614"
/codon start=1
/evidence=not experimental
/product="d407f17.1 (TAB1 (TAKI binding protein 1))"
/protein_id="CAB55304.1"
/db_xref="GI:5834565"
/translation="MAQRSLQSEQSPWTDLPCLHSLGVGSASNSRYADKGT
ESHPEDSWLKERSENCPLGVFNQGNRTYNVAORLAEILLGLNMEHADY
RRVLLAPFVVRSFLESIDALAKASQLPRGVQHOLPRYQKILRLTLER
ETSGAMAVAVLNNKLYAVNVTNRLLCKSTYDGLQVTLQNLDRHTEDELFRL
SULGIDAGKIKQVGIICQESTRIIGDKVKKGYTDIDLSAAKSPIIAPELHGQ
PLDVGTFVLVMSBGLYKALEAHGPGQNGEIAMIDTEPAKSLDVAQAVVDV
KRHSDFASGGERAPCBREHDTLLVRNFGYPLGEMSOPTSPAPAGRAVVDV
PYSSAOSTSKTSLVLPSCQMVNNGASSTIDEATPTLTNOSTNTHT
OSSSSSDGGLEFRSRPAHSIPGEDGRVPEVDFEFRLMSVDHGESVTA"
10423. .10665
/note="MIR repeat: matches 3. .261 of consensus"
10894. .11003
/note="L2 repeat: matches 2483. .2587 of consensus"
11004. .11310
/note="AluY repeat: matches 2. .308 of consensus"
11311. .11418
/note="L2 repeat: matches 2587. .2710 of consensus"
12433. .12563
/note="FLAM C repeat: matches 2. .133 of consensus"
12622. .13098
/gene="d407f17.1"
/note="match: GSS: Em:AG601543"
12823. .13100
/note="AluSg repeat: matches 1. .279 of consensus"
13182. .13227
/note="MIR repeat: matches 86. .131 of consensus"
13351. .13887
/note="L2 repeat: matches 2087. .2749 of consensus"
13888. .14192
/note="AluSg repeat: matches 1. .298 of consensus"
14193. .14515
/note="L2 repeat: matches 1790. .2087 of consensus"
14551. .14831
/note="AluO repeat: matches 1. .298 of consensus"
16401. .16671
/note="AluSg1 repeat: matches 35. .305 of consensus"
16869. .17181
/note="AluY repeat: matches 1. .310 of consensus"
17473. .17722
/note="L1 repeat: matches 4070. .4338 of consensus"
17724. .18033
/note="AluY repeat: matches 1. .308 of consensus"
18034. .18168
/note="AluY repeat: matches 1. .135 of consensus"

Alignment Scores:
Pred. No.: 3,32e-23 Length: 69660
Score: 547.50 Matches: 296
Percent Similarity: 18.93% Conservative: 44
Best Local Similarity: 16.48% Mismatches: 109
Query Match: 21.22% Indels: 1348
DB: 9 Gaps: 19
US-09-830-144-4 (1-504) x HS407F17 (1-69660)
Qy 8 LeuleuginsergiugluginInProSerTyrThrAspAspleuProleuCyHsIleuSer 27
:::|
Db 25243 GTCCCTTCCCGAGGACGACGACCGAAGCTTGACATGATGCTTGTGCGACCTCTCT
Qy 28 G1YValG1SerAlaSerAsnArgSerTyrSerAlaAspGlyThrGluSerHis 47
|
Db 25303 GGGGTGGCTCAGCTCCCAACCGGAGTACTGCTGATGGAAGGAGCATGAGAGGCAC 25362
Qy 48 ProProGluAspSerTyrPleuLysPhe----- 56
|
Db 25363 CCGCAGAGGACAGCTGCTCAAGTTTCAGGTGTGTGTGCCAGATTTTGTGTGGGCC 25422
Qy 56 ----- 56
Db 25423 GGGAGTTGGTGTGTTGCCAAGCAAGAAAGACACGACTTGCAGCTTCTGTATGG 25482
Qy 56 ----- 56
Db 25483 CTTCGACATGACATGTGCGCCGTGAGAGTGGCCTGTGCTGTCTTGGCCAAAGGCGCTG 25542
Qy 56 ----- 56
Db 25543 CTCTGATGGGGTAGCGTAGACATGGGAGGAGGATATCCGAATGTCATAGCCAGAGTG 25602
Qy 56 ----- 56
Db 25603 AAATGATGGCTAAAGACGGGGGACCCAGGAGGCCCTTGAGCTGACGCTGTGCGTT 25662
Qy 56 ----- 56
Db 25663 TAGTCTCCCAATTCTTCTTCTTCTTCTCCCTCCACCTCGTAGAGACCTGATCGAGG 25722
Qy 57 ----- 57
Db 25723 CTCCTGTGCCCTCTCCCTCTTCACAGAGAGAAACAAGCTTCTGTATGGGCTCTTC 25782
Qy 69 AsnGlyTyrAspGlyAsnArgValThrAspPheValAlaGlnArgLeuSerAlaGlu 88
|
Db 25783 AACGGCTATGATGGCAACCGATGACCAACTTCTGTGCCAGCGGCTGTCCGAGAGCTC 25842
Qy 89 LeuleuGlyInleuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuleuGln 108
|
Db 25843 CTGCTGGCCAGCTGAATGCGAGACGCGGAGCCGATGTGGCGGTGTGCTGCTGACG 25902
Qy 108 ----- 108
Db 25903 GTAATGTGTCGGGGCCACAGTAGACCCACACATCATGTCCCAACCCCAAGCTTGG 25962
Qy 108 ----- 108
Db 25963 GCCCTGACCTTAGCATGTGGCCAGGGTGTGTGAAGATCCTGCGCGGCCCTTACCC 26022
Qy 108 ----- 108
Db 26023 CAGTAGAGAGAGAGCTCCAGCGTAGAGCCCCCACCACCAAGAGTCCAGAGCACGCA 26082
Qy 108 ----- 108
Db 26083 GCGAGCAGAGGCTTCATATGCTGTGCAGAGAGTACTGTGAGACGCGGGGATTTAG 26142
Qy 108 ----- 108
Db 26143 GGATGGAGCTTGGAGAGAGAGGTGTGAGGTGGAGCGGCGCAAGGCTGTGATAATGGGG 26202

Qy 108 ----- 108
Db 26203 TCAITTAGAGTACCCCTTTCTTCTCTATGTGGTCAGGTGCTCAGCCTCCAGGTGCAGGA 26262
Qy 108 ----- 108
Db 26263 GCCATCCCTGGCGTCCAGGAAGACCTTGCTCCCTTTCTGAGGGCGCGCCCTCA 26322
Qy 108 ----- 108
Db 26323 TTGACTGTTCCACTTTCCTGCTTGGCTGTGGCAGTTTTCCTCTGGCATCTGCTTTC 26382
Qy 108 ----- 108
Db 26383 AGGAGCATGTCTCAGGCCCAFTTCAGATGAGAAGATGGCTTCTGTTCCCGAGAGGGGT 26442
Qy 108 ----- 108
Db 26443 GGTGCCAGCCTTTTCCTGCCCTTCACGACCTCAGGCTCCATTTGCCAGTGATTTCTCAGCAG 26502
Qy 108 ----- 108
Db 26503 ATCTCACACAGGGGAGAGGTGTCACTGTCAACAGTGTCTCTGGGCTGGTGGGTTTGA 26562
Qy 108 ----- 108
Db 26563 CAGAAGCCTCCAGAAAGAGTTGTAGTAATCATACGAGCTGTCTATAGGCTCGCAGTTT 26622
Qy 108 ----- 108
Db 26623 CTCTGAGCAGTTGCCCTTGCTGAGTGCCAGGTTGGACTGTCTTACCTAAACCCCTGGAACAA 26682
Qy 108 ----- 108
Db 26683 CACTTAACCTCCGTTGTGAGACTGAGAGGGGCCAGAGGTCAACACGCTGGGCTACGC 26742
Qy 108 ----- 108
Db 26743 CAAGCCTTTGCTTTAGCTTCGGCACTGTCTCTGCTGCCCTCCAGCGTCTCAGTGTGCC 26802
Qy 108 ----- 108
Db 26803 TCTCAGGGTGCACACCCCTTCGAGGCTGAAGGCTTTGTCAAAGACAITTGATCTGCAGG 26862
Qy 108 ----- 108
Db 26863 AAGCAGCGGTGCCTTGCACTGTCTGGCCAGAGCGGACTGACATGTGAAAGCTCCAT 26922
Qy 108 ----- 108
Db 26923 CACACAAGAACCTGCAGTGAAGACAGCAAGCTGCTGCTCTGATTAAATAGAGGACATTTT 26982
Qy 109 -----Ala 109
Db 26983 GGCACCACTGACAGTGGTGTTTTGAACAGCCCTTTGCCCTGTCTGTGTCCTAGGCC 27042
Qy 110 PheAspValValGluArgSerPheLeuGluSerIleAspAlaLeuAlaGluLysAla 129
Db 27043 TTCGATGTGGTGGAGAGGAGCTTCTCGAGTCCATTGACACGCTTGGCTGTGAAAGGCA 27102
Qy 130 SerLeuGlnSerGlnLeuProGlu----- 137
Db 27103 AGCCTCCAGTCGCAATTGCCAGA--GGTAATTTCCCGAGCCGACACCCAGGGAGTCAAGT 27161
Qy 137 ----- 137
Db 27162 CCAGGCCCAAGCTTTGCAAGGAGCATGACTCATCTACTTTCTTGACATTACTGGCCAGA 27221
Qy 137 ----- 137
Db 27222 GCAACAGCGGTTAGGAGCAGTTCTCTGATGGGTGACACTGGTGTGTGGCCACAGGTGAGG 27281

Qy 137 ----- 137
Db 27282 GACCTCGCTGCTCTGCTGTGGCAGCTGCATGCTAAAGCTCAGCAGTGTCTGGGCAGTGG 27341
Qy 137 ----- 137
Db 27342 TGGACATGAGGAAGGGCTTTTCCAGTCTTCTGCTCTGAATGGGTCCCCGCCCTTGTCAA 27401
Qy 137 ----- 137
Db 27402 CCACCTGCTCACATCGCCTACTGTCCACCTGCAGCGCTTACCTCTGTGTGTGAGTTCGGT 27461
Qy 137 ----- 137
Db 27462 CATCTCCAGCTTTCTGGAAGAACCTTTAGCCTGGAGCTGTAGCGAAGGCTTCATCTTGCT 27521
Qy 137 ----- 137
Db 27522 GTGCCATGACCATCTGTCTGTCTGTCCCCCGGCCCTCGGATCTTTGCTCTTAAGACA 27581
Qy 137 ----- 137
Db 27582 AAGGAGGCCGGGCATGTGTGCTCAGCCTGTAAATCCACGACCTTGGGAGCCGAGCG 27641
Qy 137 ----- 137
Db 27642 GCGGATCACAGGTCAGGAGATCAAGACCATCTCGGCTAACATGGTGAACCCCGTCTC 27701
Qy 137 ----- 137
Db 27702 TACTAAATAACAAAAATTTAGCTGGCGTGTGTGGCGCTTGTAGTCCCAGCTACT 27761
Qy 137 ----- 137
Db 27762 CTCGAGGCTGAAGCAGAGAGAATGGTGAACCTGGGAGCGAGCTTGTAGTGAGCCGAG 27821
Qy 137 ----- 137
Db 27822 ATCATGCCACTGCACCTGTAGCCTGGGGACACAGCGAGACTCCATCTCAAAAAAAGG 27881
Qy 137 ----- 137
Db 27882 AAGAGTAAAGGAGTGGAGAGGCTTAGGAAGATGGTCCAGAGTTCTGTCTGCCCTGACC 27941
Qy 138 -----GlyValProGlnHisGlnLeuProGlnTyrGlnLysIle 151
Db 27942 CTCTGTTGATGGTTGTAGGGAGTCCCTCAGCACCGAGCTGCCTCTCAGTATCAGAAGATC 28001
Qy 152 LeuGluArgLeuLysThrLeuGluArgGluIleSerGlyGlyAlaMetAlaValValAla 171
Db 28002 CTTGAGAGACTCAAGACGTTAGAGAGGGAATTTTCGGAGGGGCCATGCGCGTTGTGGCG 28061
Qy 172 ValLeuLeuAsnAsnLysLeuTyrValAlaAsnVal----- 183
Db 28062 GTCTTCTCAACAACAAGCTCTACGTGCCAAATGTGCGTGAGCCCCCTCTGTCCCAGGG 28121
Qy 183 ----- 183
Db 28122 CAGGAGGACTGGGAGAGGTCAAGCCACAGGGGTGCGTGCATTATTGACAATCTGCTTT 28181
Qy 183 ----- 183
Db 28182 CCAGACACTTCAGCACCTTTAAACCCAGGGTCTCTGTAGACCGTTGGGTATGTCCCTCTC 28241
Qy 183 ----- 183
Db 28242 CACAGTGACGCTCAGTCCCACTGGAAGGAGAAAGACGGGATGGAGACAGGTGCC 28301
Qy 183 ----- 183
Db 28302 TGGCCTTTAGTCCCTAATTTCTGCTTTTAACTCATTTCTGGACGAGTCGATCCCATTTCTGG 28361
Qy 183 ----- 183

Db 28362 GTGGCCTTGGGGCGCGCTCTTGTGAGAGGTTTACTGGCCATGATAATACCAACAGTA 28421
Qy 183 ----- 183
Db 28422 TCCCATGAAACAGCTCCGAGATGGGGCTAGGTGACAGGACATGGGGTTTGTGAAGA 28481
Qy 183 ----- 183
Db 28482 CCAGACAGGTGACAGTTTCAATGAAAGACCTGTGAGAGACCCTTGTGATGCTGCC 28541
Qy 183 ----- 183
Db 28542 TTTTTPAATCTCCAGTGAATAATTTCCACAAATATAGAACAAATAGTAGTACTATTT 28601
Qy 183 ----- 183
Db 28602 TAAATGAATTCATGGGCCATCCCAATCCAGAGTTATCACTGTGGCCAGCCTTCT 28661
Qy 183 ----- 183
Db 28662 TTGTCCCAAGTTCTCTATTTTGAACAAATGCCAGACATGATCATGTATCTGTAA 28721
Qy 183 ----- 183
Db 28722 TAATTCAGTCTGATCTAGAGAGAGGTCTGTTTATTAAGATCAATATCTATGATT 28781
Qy 183 ----- 183
Db 28782 ACACGTAAAAAGTTAAACCAATTTCTTGATACACAGCTGCTTACTGCTGACTCA 28841
Qy 183 ----- 183
Db 28842 CCCCCTTGCTGAGCCTGTTTCCCATTTTCAGTATTTCCATGTGMAATGCTGCC 28901
Qy 183 ----- 183
Db 28902 TTTTCCCTCTGCTGCTTCCCGATGCCATATTTCTCTGTGTAGTCTTTGCTTAGC 28961
Qy 184 -----
-GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGly 197
Db 28962 TGTTCACATTTCTGCCACAGGTACAAACCGTGACCTTTATGCAATCGACAGTGGG 29021
Qy 198 LeuGlnValThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArg 217
Db 29022 TTGCAGGTGACACAGCTGAACGTGACACACACAGAGAACGAGGATGAGCTTCCGT 29081
Qy 218 LeuSerGlnLeu----- 221
Db 29082 CTTTCCAGCTGGGTAGTGGGAGAGTGGGAGCGGAGCTGATCCCATGGGCTCACCC 29141
Qy 221 ----- 221
Db 29142 TTGGCTGCTTTGTGTGGGTAGAGAGCGCTGTGTAGAGGGGCTGTGATCTTGGGC 29201
Qy 221 ----- 221
Db 29202 TCCCAAGCAGCTGCTGGGTTTCATCCACAGACTGCCGCTTACTGTTGGTGTATGG 29261
Qy 221 ----- 221
Db 29262 AACTTAAGTTTCTGTGCTCAGATCCCGCGTGTAAATGAGATTAATGTAACCTACC 29321
Qy 221 ----- 221
Db 29322 TCGGAGTGTAGAGAGATTCAATTAGTGCATGTGAAATGCTTCTCGTAGAGTGCCTGCACA 29381
Qy 221 ----- 221
Db 29382 CAGAAACACTTACTGTATTGTGTGAGACTGAAGAGGCCAAAGAAGTCCAGGAGCCCC 29441
Qy 221 ----- 221

Db 29442 AGCTGCTGTAGAGTGTCTCTTCCCTTCCAGTGAAGCTCCAGGCAGATGTGGGGCACAG 29501
Qy 221 ----- 221
Db 29502 GGATGGCAGGGAGACGGCAGAGCAAGCTGCTCCGTGCCAGTGTGTCTTGAGAGACGGC 29561
Qy 221 ----- 221
Db 29562 TCTGAGTAGGTGTGTCTGTATGGGCGTGGGACTAGAGACACAGGACTTTGGGTGAG 29621
Qy 221 ----- 221
Db 29622 CTGCTCCAGAGTCCCTGTGTGTGTCTTATTCCTATTCAGTGTGCTTATTGCT 29681
Qy 221 ----- 221
Db 29682 TCTTCCATGACTGTGTCTGTCTGCCCTTGTGTTGTTGCTCTTTGTGAACAAGAGCAG 29741
Qy 222 -----
-GlyLeuAspAlaGlyLysIleLysGlnValGlyIle 223
Db 29742 GATTGTGCACCTGTTTCCCTCCGTAGGCTTGATGCTGMAAGATCAAGCAGGTGGGAT 29801
Qy 223 eileCysGlyGlnGluSerThrArgArgIleGlyAspTyrLysValLysTyrGlyTyrTh 253
Db 29802 CATCTGTGGCAGAGAGCACCAGCGGATCGGGATTCAGAGTTAAATATGCTACAC 29861
Qy 253 AspIleAspLeuLeuSer-----AlaAlaLysSerLysProIleIleAlaGluProG 271
Db 29862 GGACATTGACCTTCTCAGTATAGTGCAGCCAGCTGTCCCTGTGCTTCAAAAGACAGA 29921
Qy 271 uileHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluG 291
Db 29922 A-----GGTCTTGAAGGAGCCAAAGATGAGATTAATGTGCTG-----GG 29963
Qy 291 Y-LeuTyrLysAlaLeuGlnAlaAlaHisGly-----ProGlyGlnAlaAsnG 307
Db 29964 AGGTGCAGGCTGCAGTACGCCGTGATCATGCTGCAGCTCCAGCTTGGCCAAACAGCA 30023
Qy 307 IngIuIleAla-----AlaMetIleAspThrGluPheAlaLysGlnThrSerLeuA 324
Db 30024 AGACCTGTCTCAAAAAAAGAGTCCAGAAAGCTGGGCGACAGCG---AGACTGT 30080
Qy 324 spAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAspThrPheAlaS 344
Db 30081 AGGCCCAAGTGCACATGTGCAGTCAAAAGTTCACAGGCCAGCAAACTCTTCTTCC 30140
Qy 344 ergIlyGlyGluArg-----AlaArgPheCysProArgHisGluAspMetThrLeuL 361
Db 30141 TC---CCAGAGAGTGTGTGTGGGGCAGGTTAAGCCCAAGCAGGCGCAGAGTGGTA 30197
Qy 361 euValArgAsnPheGlyTyrProLeuGlyGluMet-----SerGlnProThrPro 377
Db 30198 TAGAAACAAGGGGTGAGTGTAGCTGGCCAAAGCTAGAGGCGACAACTAGTCCACAAG 30257
Qy 378 -----SerProAlaProAlaAlaGlyArgValTyrProValSer-----V 392
Db 30258 GAGGCAAGCCGAAAGCCAGAGTGGCATGAGGATGAGACAGACAGCCCTGTGTGCTGCT 30317
Qy 392 alProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetP 412
Db 30318 GCCCATGGCTGCAGCAGAG-----TGTGAGATGAAATGTGTATTTC 30359
Qy 412 roSerGln-----GlyIleMetValAsnG 420
Db 30360 CCTCTGAGAGACCTCTCTGAGACATCACCCAGCGCTCATGGCAGAGAAATGGGACGTC 30419
Qy 420 lylAlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProT 440
Db 30420 ATCATCAACAGGCCAAGCCCGTCAGAGCTGTGAGATGATCTTCTTTATGACTTCCCTA 30479
Qy 440 hrlLeuThrLeuGlnSerThrAsnThrHis-ThrGlnSerSerSerSerSerAspGly 459
Db 30480 CTCAGGCCCTGTAGGAGATCATCTGCATCATCCCCAGAGGCAATTTTCAATAGAGAAACT 30539

QY 460 GlyLeuPheArgSerArgProAlaHisSerLeuProGly 473
 Db 30540 GAGCATGGAGATCATCAGGACGCCAC-----CCTGGT 30572
 RESULT 13
 AC127784/C
 LOCUS
 DEFINITION
 AC127784 176665 bp DNA linear HTG 19-JUL-2002
 Rattus norvegicus clone CH230-131B6, *** SEQUENCING IN PROGRESS
 ***, 59 unordered pieces.
 AC127784
 AC127784.1 GI:21908163
 HTG; HTGS_PHASE1.
 Rattus norvegicus
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 176665)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbaria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louissegh,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okwuonu,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Umami,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G., and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 176665)
 Worley,K.C.
 Direct Submission
 Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information

Center project name: GZWK
 Center clone name: CH230-131B6
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 120196 bases at least Q40
 Consensus quality: 127192 bases at least Q30
 Consensus quality: 131862 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 59 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1
 1032: contig of 1032 bp in length
 1132: gap of unknown length
 2439: contig of 1307 bp in length
 2539: gap of unknown length
 2440
 3610: contig of 1071 bp in length
 3710: gap of unknown length
 4955: contig of 1245 bp in length
 4956
 5055: gap of unknown length
 7048: contig of 1993 bp in length
 7148: gap of unknown length
 8902: contig of 1754 bp in length
 8903
 10567: contig of 1565 bp in length
 10568
 12258: contig of 1591 bp in length
 12358: gap of unknown length
 14216: contig of 1858 bp in length
 14316: gap of unknown length
 14317
 15760: contig of 1444 bp in length
 15761
 15860: gap of unknown length
 16935: contig of 1075 bp in length
 17035: gap of unknown length
 18984: contig of 1949 bp in length
 18985
 20295: contig of 1211 bp in length
 20296
 20395: gap of unknown length
 21672: contig of 1277 bp in length
 21673
 21772: gap of unknown length
 23061: contig of 1289 bp in length
 23161: gap of unknown length
 24596: contig of 1435 bp in length
 24597
 24696: gap of unknown length
 26381: contig of 1685 bp in length
 26382
 26481: gap of unknown length
 29059: contig of 2578 bp in length
 29159: gap of unknown length
 30322: contig of 1163 bp in length
 30422: gap of unknown length
 32594: contig of 2172 bp in length
 32595
 32694: gap of unknown length
 3702: contig of 2008 bp in length
 34802: gap of unknown length
 37016: contig of 2214 bp in length
 37116: gap of unknown length
 37117
 38321: contig of 1205 bp in length
 38421: gap of unknown length
 38422
 40849: contig of 2428 bp in length
 40850
 40949: gap of unknown length
 43133: contig of 2184 bp in length
 43233: gap of unknown length
 43134
 45386: contig of 2153 bp in length
 45487: gap of unknown length
 45487
 48376: contig of 2890 bp in length

```
* 48377 48476: gap of unknown length
* 48477 51407: contig of 2931 bp in length
* 51408 51507: gap of unknown length
* 51508 53492: contig of 1985 bp in length
* 53493 53592: gap of unknown length
* 53593 55777: contig of 2185 bp in length
* 55778 55877: gap of unknown length
* 55878 58753: contig of 2776 bp in length
* 58754 61259: contig of 2506 bp in length
* 61260 61359: gap of unknown length
* 61360 63193: contig of 1834 bp in length
* 63194 63294: gap of unknown length
* 63294 66481: contig of 3188 bp in length
* 66482 66581: gap of unknown length
* 66582 68046: contig of 1465 bp in length
* 68047 68146: gap of unknown length
* 68147 72431: contig of 4285 bp in length
* 72432 72531: gap of unknown length
* 72532 74844: contig of 2313 bp in length
* 74845 74944: gap of unknown length
* 74945 77951: contig of 3007 bp in length
* 77952 78051: gap of unknown length
* 78052 81664: contig of 3613 bp in length
* 81665 84688: gap of unknown length
* 84689 84788: contig of 2924 bp in length
* 84789 87987: contig of 3199 bp in length
* 87988 88087: gap of unknown length
* 88088 91325: contig of 3238 bp in length
* 91326 91425: gap of unknown length
* 91426 94907: contig of 3482 bp in length
* 94908 95007: gap of unknown length
* 95008 98189: contig of 3182 bp in length
* 98190 98289: gap of unknown length
* 98290 102290: contig of 4001 bp in length
* 102291 102390: gap of unknown length
* 102391 105219: contig of 2829 bp in length
* 105220 105319: gap of unknown length
* 105320 108533: contig of 3214 bp in length
* 108534 108633: gap of unknown length
* 112138 112138: contig of 3505 bp in length
* 112139 112238: gap of unknown length
* 112239 116139: contig of 3901 bp in length
* 116140 116239: gap of unknown length
* 116240 121227: contig of 4988 bp in length
* 121228 121327: gap of unknown length
* 121328 125446: contig of 4119 bp in length
* 125447 125546: gap of unknown length
* 125547 128969: contig of 3423 bp in length
* 128970 129069: gap of unknown length
* 129070 133782: contig of 4713 bp in length
* 133783 133882: gap of unknown length
* 133883 141065: contig of 7183 bp in length
* 141066 141165: gap of unknown length
* 141166 147195: contig of 6029 bp in length
* 147195 147294: gap of unknown length
* 147295 154395: contig of 7101 bp in length
* 154396 154495: gap of unknown length
* 154496 160801: contig of 6306 bp in length
* 160802 160901: gap of unknown length
```

Alignment Scores:

```
Pred. No.: 4.32e-17 Length: 176665
Score: 457.00 Matches: 158
Percent Similarity: 23.14% Conservative: 4
Best Local Similarity: 22.57% Mismatches: 11
Query Match: 17.71% Indels: 528
DB: 2 Gaps: 3
```

US-09-830-144-4 (1-504) X AC127784 (1-176665)

QY 12 GUGINGInProSeTrrhPnAspAspLeuProLeuCysHisLeu-SerGIyValGIySe 31

```
||||:|||||
Db 124370 GAGGAGCAGCCAGCTGGAGAGATGACCTGCTGCTCTCGCGGGTGGTTC 124311
QY 31 rAlaSerAsnArgSerTySerAlaAspGIyVsgIyThnIuSerHiSProProGIuAs 51
Db 124310 AGCTTCGGGGGAGGAGTACTCTGTATGSCAG-GGCACGGAGAGTACCTCTCGSAGGA 124252
QY 51 pSerTrrPLeu----- 54
Db 124251 CAACTGGCTAGTTTCAGGTAAAGTGGGTGGCTGAGCANNAAGATCATGACGAAG 124192
QY 54 ----- 54
Db 124191 AAGATCTGCTCTGCTACTGATGTATGAGCTTGCTGTGATGAGGCCCATGGAA 124132
QY 54 ----- 54
Db 124131 GTGTGCACTGTGTGGGACAGGGGTGGACGAGCTGGCAGGAAGATTCGTGTGAGATG 124072
QY 54 ----- 54
Db 124071 ATGGACCCAGTGCAGGGTCCAGATGTGTGCTGAACCTTCTGTGTATGCTCCCTCC 124012
QY 54 ----- 54
Db 124011 CCAATTCTTCTCTTCTTCTCTTCAACACTGCATGCCCCCTGTTAGCCCTCCCTCC 123952
QY 55 -----LysPheArgSerGIyAsnAsnCySPhLeuTyGIyValPheAsnGIyTYTA 72
Db 123951 ACCCTCCCTGTGGTGAAGAGGAAACACTGCTTCTGTATGGGGTCTTCAATGGCTTAG 123892
QY 72 spGIyAsnArgValThnAsnPhyValAlaGIyArgLeuSerAlaGIyLeuLeuGIyG 92
Db 123891 ATGGCAACCCGCTGACCAACTTGTGTGGCGCAGAGACTCTGCAGAGCTCTGCTGGGCC 123832
QY 92 InLeuAsnAlaGIyHisAlaGIyValAspValArgArgValLeu----- 106
Db 123831 AGCTCAACACCGAACAACACAGATGTGTGTGCGAGGGTCTGTGTCAGATGAGGAT 123772
QY 106 ----- 106
Db 123771 TGGGCTAGTGCACAACAGCTGACTGCTGTGTCTTCACTCCCTGGAACTTGGGCC 123712
QY 106 ----- 106
Db 123711 CTGCATCTGTAGTGTAGACAGGCTGATCTGAAGCTGCTGTCTTCACTCCCTAGCA 123652
QY 106 ----- 106
Db 123651 AGATCATTTTGGGGTACACTTCATGTGTGGCAGGGGCTCTGGCCACAGAGATCCAGGAC 123592
QY 106 ----- 106
Db 123591 AGGCAGGAGACAGAGACTTTTCATTGTGTAGACAGACTGACTTGGAGAAAGGAGACTG 123532
QY 106 ----- 106
Db 123531 GGAACAGAGCTGTGTAGATAGATACCTTTGTTGAGATTCTTATTTCTTCTTAC 123472
QY 106 ----- 106
Db 123471 CTAGGAAGTTTCTTGTGTCTTGTAGAAATTGGCCCTAGATTCTTCTGTGTGTC 123412
QY 106 ----- 106
Db 123411 TCAGTCTTGTAGACTGAGCTGTGCAGAGGCTAGACAGAGAAAGCTTCTGTCTTAC 123352
QY 106 ----- 106
Db 123351 CAGTGTCAACAGCTCTGTATCAAGGCCGAGGTTGTATGCCCCCTACAGACTCAGTAC 123292
QY 106 ----- 106
```

Db 123291	CAAGCACTCTCGACTGTCTCACTTCACATGAACCCAAATGTGGAGCACTGCTGTTCA	123232
QY 106	-----	106
Db 123231	CCCAACACTCTTGAGCGCAGGGCTGGCCCTTCTCTCTCATTTCTTGAGGCTGCATAA	123172
QY 106	-----	106
Db 123171	GTAGCAACCTTTTCCCGTCCACAGGTAAGGTCTTTGTCAAACAGTGGACAGAAAGCA	123112
QY 106	-----	106
Db 123111	CACATATGAGGCAGAAATGATATGTCACTGTCAACAGAGCCATCCATGAATGCCATGG	123052
QY 106	-----	106
Db 123051	ATCTGTTACTCTCCATAGGAGACATCTTAACAGGAGGTGATGGGTATTTTGTAGCCCAACC	122992
QY 107	-----LeuGlnAlaPheAspValValGluArgSerPheLeuGluSer	120
Db 122991	TTGCCTTGCCCTGTGTCTCTTCAGGCGCTTCGATGTGGTGAGAGGAGCTTCTCTGGAGTCT	122932
QY 121	IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGlu	137
Db 122931	ATCGATGATGCTTGGCTGAGAAGCAAGCCTCCAGTCCAGCTGCCGAGGATACCGAC	122872
QY 137	-----	137
Db 122871	TGTCGTACTAGTACTGCTGGGTGCCAGCACTCACCTCTGTATCATCTTGCACCCACAGG	122812
QY 137	-----	137
Db 122811	CACAGTAAGGCTGCAGCAGGAATCATGAGCAGCATCTGTGGCAGGCTGGGTGGC	122752
QY 137	-----	137
Db 122751	ACTGGGTATGTCCCTGTAGAGACTCCAGCCACAGTGGGCAGCTTTTCTCGAGGCTCA	122692
QY 137	-----	137
Db 122691	TCCTTAGTGGCAGGATAGGTGGAAGAAAGGCTGTCCAGGTTCTGTGATCTGTTTCA	122632
QY 137	-----	137
Db 122631	TCCAGAGCCAGGCTCACCTTCAGGCTGCCCTAGTTAAAGCAACAGTACCTTCCACGG	122572
QY 137	-----	137
Db 122571	CAGCCTACAGGTATCCAGCGTGGCCACACCTATGTGTCTGTCTTGGCTGCTCAA	122512
QY 137	-----	137
Db 122511	GGGCCTTGATATCTTTATCTAAAGGAAAGAGAGGATGGTAAGTGGTACTCAAGCCCCC	122452
QY 138	-----GlyValProGlnHisGlnLeuProPro	146
Db 122451	ATGATGCTCTGATCTCTGTGATGATGTTTTCAGGGGTGTCCTCCCAACACAGTGCTACT	122392
QY 147	GlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGluIleSerGlyGlyAla	166
Db 122391	CAGTATCAGAAGATCTCTGGAGAGACTCAAGGCACCTGGAGAAGGAGATCTCGGAGGGCC	122332
QY 167	MetAlaValValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGly	184
Db 122331	ATGGCTGTTGTGGCGGTCTCTTCTCAACAACAAGCTCTACGTTGCCAATGTTGT	122278
RESULT 14		
AC127784	176665 bp DNA linear HTG 19-JUL-2002	
LOCUS	AC127784	
DEFINITION	Rattus norvegicus clone CH230-131B6, *** SEQUENCING IN PROGRESS	
ACCESSION	AC127784	
VERSION	AC127784.1 GI:21908163	

HTG; HTGS_PHASE1.
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 176665)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Binage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louseged,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhiney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaik,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 176665)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information

Center project name: GZWK
Center clone name: CH230-131B6
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 120196 bases at least Q40
Consensus quality: 127192 bases at least Q30
Consensus quality: 131862 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

NOTE: This is a 'working draft' sequence. It currently consists of 59 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1	1032: contig of 1032 bp in length
1033	1132: gap of unknown length
1133	2439: contig of 1307 bp in length
2440	2539: gap of unknown length
2540	3510: contig of 1071 bp in length
3611	3711: gap of unknown length
3711	4955: contig of 1245 bp in length
4956	5055: gap of unknown length
5056	7048: contig of 1993 bp in length
7049	7148: gap of unknown length
7149	8902: contig of 1754 bp in length
8903	9002: gap of unknown length
9003	10567: contig of 1565 bp in length
10568	10667: gap of unknown length
10668	12258: contig of 1591 bp in length
12259	12258: gap of unknown length
12359	14216: contig of 1858 bp in length
14217	14316: gap of unknown length
14317	15760: contig of 1444 bp in length
15761	15860: gap of unknown length
15861	16935: contig of 1075 bp in length
16936	17035: gap of unknown length
17036	18984: contig of 1949 bp in length
18985	19084: gap of unknown length
19085	20295: contig of 1211 bp in length
20296	20395: gap of unknown length
20396	21672: contig of 1277 bp in length
21673	21772: gap of unknown length
21773	23061: contig of 1289 bp in length
23062	23161: gap of unknown length
23162	24596: contig of 1435 bp in length
24597	24696: gap of unknown length
24697	26381: contig of 1685 bp in length
26382	26881: gap of unknown length
26882	29059: contig of 2578 bp in length
29060	29159: gap of unknown length
29160	30322: contig of 1163 bp in length
30323	30422: gap of unknown length
30423	32594: contig of 2172 bp in length
32595	32694: gap of unknown length
32695	34702: contig of 2008 bp in length
34703	34802: gap of unknown length
34803	37016: contig of 2214 bp in length
37017	37116: gap of unknown length
37117	38921: contig of 1205 bp in length
38922	38921: gap of unknown length
38922	40849: contig of 2428 bp in length
40850	40949: gap of unknown length
40950	43333: contig of 2184 bp in length
43134	43333: gap of unknown length
43234	45886: contig of 2153 bp in length
45887	45886: gap of unknown length
45887	48376: contig of 2890 bp in length
48377	48476: gap of unknown length
48477	51007: contig of 2931 bp in length
51408	51507: gap of unknown length
51508	53992: contig of 1985 bp in length
53493	53592: gap of unknown length
53593	55777: contig of 2185 bp in length
55778	55877: gap of unknown length
55878	58653: contig of 2776 bp in length
58654	58753: gap of unknown length
58754	61593: contig of 2506 bp in length
61260	61359: gap of unknown length
61360	61393: contig of 1834 bp in length

631394	63293: gap of unknown length
633994	66481: contig of 3188 bp in length
66581	66581: gap of unknown length
66582	68046: contig of 1465 bp in length
68047	68146: gap of unknown length
68147	72431: contig of 4285 bp in length
72432	72531: gap of unknown length
72532	74844: contig of 2133 bp in length
74845	74944: gap of unknown length
74945	77951: contig of 3007 bp in length
77952	78051: gap of unknown length
78052	81654: contig of 3613 bp in length
81655	81764: gap of unknown length
81765	84688: contig of 2924 bp in length
84689	84788: gap of unknown length
84789	87987: contig of 3199 bp in length
87988	88087: gap of unknown length
88088	91325: contig of 3238 bp in length
91326	91425: gap of unknown length
91426	94907: contig of 3442 bp in length
94908	95007: gap of unknown length
95008	98189: contig of 3182 bp in length
98190	98289: gap of unknown length
98290	102290: contig of 4001 bp in length
102291	102390: gap of unknown length
102391	105219: contig of 2829 bp in length
105220	105319: gap of unknown length
105320	108533: contig of 3214 bp in length
108534	108633: gap of unknown length
108634	112138: contig of 3505 bp in length
112139	112238: gap of unknown length
112239	116139: contig of 3901 bp in length
116140	116239: gap of unknown length
116240	121227: contig of 4988 bp in length
121228	121327: gap of unknown length
121328	125446: contig of 4119 bp in length
125447	125546: gap of unknown length
125547	128969: contig of 3423 bp in length
128970	129069: gap of unknown length
129070	133782: contig of 4713 bp in length
133783	133882: gap of unknown length
133883	141065: contig of 7183 bp in length
141066	141165: gap of unknown length
141166	147194: contig of 6029 bp in length
147195	147294: gap of unknown length
147295	154395: contig of 7101 bp in length
154396	154495: gap of unknown length
154496	160801: contig of 6306 bp in length
160802	160901: gap of unknown length

Alignment Scores:		Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
US-03-830-144-4 (1-504)	x AC127784 (1-176665)	1.98e-14	414.00	25.17%	22.90%	16.05%	2
		Length:	176665	Matches:	13	Mismatches:	20
						Indels:	408
						Gaps:	4
QY	335	LysArgIleHisSerAspThrPheAlaSerGlyGluArgAlaArgPheCysProArg	354				
Db	161718	AAGAGACTCTCAAGGAGACATCTACCTGGGC-----TGTCCCAAG	161759				
QY	355	His-----GluAspMetThrLeuValArgAsn-----	364				
Db	161760	AGTGAGGGCCACAGCTGAGACTTGAGAGATTAACCTTTACTACGAGTAGCTTGA	161819				
QY	365	---PheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla-ProAlaAl	383				
Db	161820	GCCCACTACTCTTGCTTTCCTCCCAACCAATTTCTTCCCTGAGCCCTCTTCTTGG	161879				
QY	383	agIcVlVArvAlTvrProValSerValProTyrSerSerAlaGluSerThrSerIvsth	403				


```
|||||
Db 161880 AGGGCGCGTGTACCCCTGTTCTGTGCCCCCTACTCAAGTGCACGAGCACCACGAAGAC 161939
QY 403 rSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGlyAlaHisSe 423
|||||
Db 161940 CAGTGTGACTCTGTCCCTCGTCTCATGCTTCTCAGGGCCAGATGGTCAAAGGCTCTCACAG 161999
QY 423 rAlaSerThrLeuAspGluAlaThrProThrLeuThr----- 435
|||||
Db 162000 TGCCTCCACCTTGGATGAAGCCACTCTACTCTCACTAAGTAAGCTATCCTACTCTAGGT 162059
QY 435 ----- 435
Db 162060 ACCTGGAGAGGGACAAACCCCTGTACAGCAAGCCCTAGGGCTCAGTAGGGTGGGTAAGA 162119
QY 435 ----- 435
Db 162120 GGCTGGGGTGTGGGACCCTGGGGTCTTTGACCTCCAATTCCTCCCTTCCCTTTGGGTA 162179
QY 435 ----- 435
Db 162180 GGGAGTCTTATTAGACACAGGTAGGAAGGACAACTACTGTTAATTCTAGTAGTGTGA 162239
QY 435 ----- 435
Db 162240 CAGGACCTGGGGCAAGTGGAGGTTTCAGGGCAAGGTGGACAAGGGGCCCGAGCAAGGAGC 162299
QY 435 ----- 435
Db 162300 AAGTGGCTGACTGATGTCCTGGGGCAGCTTTCATCCATGGCAGCTTCTCTGGTTTGCTC 162359
QY 435 ----- 435
Db 162360 CTTCTGTTTTGGAGCTGTAAACAGGGTCACTGGCTTCACGAATCTGTACTTGGCCAGATT 162419
QY 435 ----- 435
Db 162420 CATTCCTGGCCTGTACTGGGCTTTCAGAGCTGGGCTCCCTCACTGTGACGGACCATTTG 162479
QY 435 ----- 435
Db 162480 CTGTCACTTGTCTAGTTTAAAGGACGAGGAGGTAGAGGCTATGGGCTCTGAGAGGGGCA 162539
QY 435 ----- 435
Db 162540 TGGTGCAGACCTGCTGCACAGGCTATTATCTACCGCCCCCTTCATTTCTGTGACTGTGT 162599
QY 435 ----- 435
Db 162600 ACTCTGCCCCCACCACCTCTCTCAGGGGATCTGGTTTCCACATTCACAGCCCTGCACGTG 162659
QY 435 ----- 435
Db 162660 GTGGAATAATACAGCTGCCCCCACCAGCCCTCTTCAGGATAGACTTAGCCGATGTGG 162719
QY 435 ----- 435
Db 162720 CCAATTAGAGGGCTGGCAGGAGGAGTGCCTGTTTCTCTGCTCTCCCTTTCAGAGGGC 162779
QY 435 ----- 435
Db 162780 CTCCCATCTGTCCCAAGTCTTTGCTGCTGCTTTTCCACCACCACCTCTCTGGATAGA 162839
QY 435 ----- 435
Db 162840 GGGCCTGGTGTAGACATCTGGGCATCATGGGTAGCGGTGTGGTGTCCCCCAGGACAGTCAA 162899
QY 435 ----- 435
Db 162900 CATAGCTCCCCCAACCCCCCACCCTGCTTTCCCGAGGAACCGCCAGCTGTGTCCACACA 162959
QY 435 ----- 435
```

```
Db 162960 GCCAACCCAGGAAAATTTCATGCTTCTGTGTGACTGCTGCTGAACCTCACAGGTCTCTTCC 163019
QY 435 ----- 435
Db 163020 TCTGGCTCTGAGCCTTATTCTGTACCTAGTCAAGCTGTCTTTTAAACCATCTCACCC 163079
QY 435 ----- 435
Db 163080 CACTGGCTTTTATGATATTTGGCAGAGAGGCCTGAGCCTAAAGGGCCACTCTTTTCACT 163139
QY 435 ----- 435
Db 163140 GGTGGTGGGAAGTTGACGGCCCTTCAGGCAGCTACCCAGACTGATTGCTTCCCTCTCT 163199
QY 436 -----AsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerS 454
|||||
Db 163200 TCTGTAGCCAGAGCCCACTCTGACCTGACGCTACCAACACCACACCCAGAGGACGA 163259
QY 454 erSerSerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProProGlyG 474
|||||
Db 163260 GCTCCAGCTCGACGGGGCCCTTTCGCTCCAGACCGGCTCACTCACTTCCACCTGGAG 163319
QY 474 luAspGlyArgValGluProTyrValAspPheAlaGluPheTyrArgLeuTrpSerVala 494
Db 163320 AGGATGGCAGGGTCCAGCCCTATGTGAGCTTTGTGAGTTTACCGACTCTGGAGCGTGG 163379
QY 494 spHisGlyGluGlnSerValValThrAlaPro 504
|||||
Db 163380 ACCATGGCAGCAGGTGTGATACGGCACCT 163411
RESULT 15
LOCUS AC127924 141498 bp DNA linear HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-23318, *** SEQUENCING IN PROGRESS
***, 68 unordered pieces.
AC127924
AC127924.1 GI:21908451
VERSION HTG: HTGS PHASE1.
KEYWORDS Rattus norvegicus.
SOURCE Rattus norvegicus.
ORGANISM Rattus norvegicus.
REFERENCE 1 (bases 1 to 141498)
AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbaria,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jollivet,S., Joudan,S.,
Karleson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,C., Liu,J., Liu,W., Louleghed,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Peters, L., Pickens, R., Primus, E., Pu, L., Qulles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojoubokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinsan, R., Wang, O.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Unpublished
Direct Submission
2 (bases 1 to 141498)
Worley, K.C.
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center of Medicine
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: KXAS
Center clone name: CH230-23318
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 86446 bases at least Q40
Consensus quality: 92843 bases at least Q30
Consensus quality: 96514 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 68 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1030: contig of 1030 bp in length
1031: gap of unknown length
1131: contig of 1132 bp in length
2263: gap of unknown length
2362: gap of unknown length
3440: contig of 1078 bp in length
3441: gap of unknown length
3540: gap of unknown length
5167: contig of 1627 bp in length
5168: gap of unknown length
5267: gap of unknown length
5268: contig of 1113 bp in length
6581: gap of unknown length
6680: gap of unknown length
6681: contig of 1203 bp in length
7884: gap of unknown length
7984: contig of 1038 bp in length
9021: gap of unknown length
9022: gap of unknown length
9122: contig of 1704 bp in length
10825: gap of unknown length
10826: gap of unknown length
10926: contig of 1158 bp in length
12083: gap of unknown length
12084: gap of unknown length
12183: gap of unknown length
12184: contig of 1127 bp in length
13111: gap of unknown length
13410: gap of unknown length
13411: gap of unknown length
14464: contig of 1054 bp in length
14465: gap of unknown length
14565: gap of unknown length
16304: contig of 1740 bp in length
16305: gap of unknown length
16404: gap of unknown length
16405: contig of 1935 bp in length
18339: gap of unknown length
18340: contig of 1054 bp in length
18439: gap of unknown length
18440: contig of 1054 bp in length
19493: gap of unknown length
19494: gap of unknown length
19593: gap of unknown length

19594: contig of 1705 bp in length
21298: gap of unknown length
21399: gap of unknown length
21399: contig of 1139 bp in length
22537: gap of unknown length
22538: gap of unknown length
22637: gap of unknown length
24129: contig of 1492 bp in length
24130: gap of unknown length
24230: gap of unknown length
24230: contig of 1048 bp in length
25278: gap of unknown length
25278: gap of unknown length
25378: contig of 1701 bp in length
27079: gap of unknown length
27179: gap of unknown length
29508: contig of 2330 bp in length
29509: gap of unknown length
29609: contig of 1812 bp in length
31421: gap of unknown length
31421: gap of unknown length
31521: contig of 1714 bp in length
33234: gap of unknown length
33235: gap of unknown length
33335: contig of 1776 bp in length
35110: gap of unknown length
35111: gap of unknown length
35210: gap of unknown length
35211: contig of 1465 bp in length
36576: gap of unknown length
36576: gap of unknown length
36776: contig of 1119 bp in length
37894: gap of unknown length
37895: gap of unknown length
37995: gap of unknown length
39970: contig of 1976 bp in length
40070: gap of unknown length
40071: gap of unknown length
41623: contig of 1553 bp in length
41624: gap of unknown length
41724: gap of unknown length
43157: contig of 1434 bp in length
43158: gap of unknown length
43257: gap of unknown length
45044: contig of 1787 bp in length
45045: gap of unknown length
45145: gap of unknown length
46505: contig of 1460 bp in length
46704: gap of unknown length
46704: gap of unknown length
46705: gap of unknown length
47898: contig of 1194 bp in length
47899: gap of unknown length
47999: gap of unknown length
49208: contig of 1210 bp in length
49309: gap of unknown length
49309: gap of unknown length
50615: contig of 1307 bp in length
50715: gap of unknown length
50716: gap of unknown length
50716: contig of 1784 bp in length
52499: gap of unknown length
52500: gap of unknown length
52500: gap of unknown length
54003: contig of 1404 bp in length
54004: gap of unknown length
54103: gap of unknown length
54104: gap of unknown length
55661: contig of 1558 bp in length
55662: gap of unknown length
55761: gap of unknown length
57610: gap of unknown length
58110: contig of 2349 bp in length
58111: gap of unknown length
61322: gap of unknown length
61322: contig of 3112 bp in length
61423: gap of unknown length
61423: gap of unknown length
62808: contig of 1366 bp in length
62809: gap of unknown length
62908: gap of unknown length
62909: gap of unknown length
65459: contig of 2551 bp in length
65460: gap of unknown length
65559: gap of unknown length
65660: gap of unknown length
66882: contig of 1323 bp in length
66883: gap of unknown length
66983: gap of unknown length
68456: contig of 1474 bp in length
68457: gap of unknown length
68556: gap of unknown length
70789: gap of unknown length
70789: contig of 2233 bp in length
70889: gap of unknown length
73103: gap of unknown length
73103: contig of 2214 bp in length
73203: gap of unknown length
73204: gap of unknown length
73204: contig of 1518 bp in length
74721: gap of unknown length
74722: gap of unknown length
74821: gap of unknown length
77416: contig of 2595 bp in length
77417: gap of unknown length
77516: gap of unknown length
77517: gap of unknown length
78949: contig of 1433 bp in length
78949: gap of unknown length
79049: gap of unknown length
81501: contig of 2452 bp in length
81501: gap of unknown length
81601: gap of unknown length
81602: gap of unknown length
83009: contig of 1408 bp in length
83109: gap of unknown length
83110: gap of unknown length
84773: contig of 1664 bp in length
84774: gap of unknown length
84873: gap of unknown length
84874: contig of 2018 bp in length

* 86892 86991: gap of unknown length
* 86992 88782: contig of 1791 bp in length
* 88783 88882: gap of unknown length
* 88883 90589: contig of 1707 bp in length
* 90590 90689: gap of unknown length
* 90690 92680: contig of 1991 bp in length
* 92681 92780: gap of unknown length
* 92781 94508: contig of 1728 bp in length
* 94509 94608: gap of unknown length
* 94609 97262: contig of 2654 bp in length
* 97263 97362: gap of unknown length
* 97363 100119: contig of 2757 bp in length
* 100120 100219: gap of unknown length

Alignment Scores:

Pred. No.: 7,03e-13 Length: 141498
Score: 387.00 Matches: 121
Percent Similarity: 45.95% Conservative: 38
Best Local Similarity: 34.97% Mismatches: 97
Query Match: 15.00% Indels: 90
DB: 2 Gaps: 15

US-09-830-144-4 (1-504) x AC127924 (1-141498)

QY 211 AsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIleLysGln 230
Db 83252 TCTAGGACAGCGCTGTGAGACTTGGAGGCTGTGGAGCTGGAGTTCGGAAGAGCCAGAT 83311
QY 231 ValGlyIleIleCysGlyGln-----GluSerThrArgArgIleGlyAsp 245
Db 83312 ACAAGATTATCTGCCCGCCAGCGTGTGCATGGTGTGCACCTTACAGCTCAAGCCGCTCT 83371
QY 246 TyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSer---Lys 264
Db 83372 TTCTTTCTAGA-----ACTGATTGGCTCAGGGAACAATTGCTAGTCTGTAGT 83422
QY 265 ProIleIleAlaGluPro-----GluIleHisGlyAlaGlnProLeuAspGlyValThr 282
Db 83423 CAGGCATTGGCTCTGCCTGTCTAGGACAGCTGCAATTCCCTCC----- 83464
QY 283 GlyPheLeuValLeuMetSerGluGlyLeuTyr-----LysAlaLeuGluAla 298
Db 83465 ---TGGGTCACTTACCTCGCCTCTCGCTCTATCTCTCAGGATCTCGATCTCTAAGTCTC 83521
QY 299 AlaHis-----GlyPro 302
Db 83522 CTACACAGGCTTTGGATCCCCCTTACCCGCCCCCTTCTCCTCATCAGGTTCT 83581
QY 303 GlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSer 322
Db 83582 CCCTCCACACCCAGGAGGTTCGCAATGATTGACACCGAATTTGCCAAGCAGACCTCC 83641
QY 323 LeuAspAlaValAlaGlnAlaValAspArgValLysArgIleHisSerAspThrPhe 342
Db 83642 CTCGATGACGTTCGCCAGGCTGTGTAGACCGTGTAAAGCGTATCCACAGTGCACACCTTT 83701
QY 343 AlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeuLeuVal 362
Db 83702 GCCAGCGGTGGGAGCGGGCCAAAGTTCTGCCACGGCATGAAGACATGACCTCTGTGGTG 83761
QY 363 ArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAlaProAla 382
Db 83762 AGGAACCTTTGGCTATCATTTAGGTGAATGAGCCAGCCGACACCCACCCGCGCCAGGT 83821
QY 383 -----AlaGlyGlyArg 386
Db 83822 ACGTATGCAATGTACTGTACAGGTAGAAGCACCTGAGATGAGTCTGGGTGGGGCAGC 83881
QY 387 Val-----TyrProValSerVal 392
Db 83882 ACAACATGAGCTGACAAACATGAGCTCGCTGTTGGAATGCAGAGGAGCGCCTGG 83941
QY 393 ProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetPro 412

Db 83942 ---CTTTGCTCAACACAGGCTGACTCTGAGCTGCAAAATCTTTATTCGTTTCTG--- 83992
QY 413 SerGlnGlyGlnMetVal-----AsnGlyAlaHis 422
Db 83993 AAAGAAGGATCAATTGCTCTCTGGCCACCCTGTTATTCTTGTGTCATGAGCACATCAGTCAC 84052
QY 423 -----SerAlaSerThrLeuAspGluAlaThrPro---ThrLeuThr 435
Db 84053 AAAGCATGGTAGAGTGCACAGCTGCCACCATCCATCGATGGTGGAGTCCCTGTTAGCCTATGG 84112
QY 436 AsnGlnSerProThrLeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSer 455
Db 84113 TCAGTGTGGTCACTTCTCACCTTTTACTAGCAACCCCTAACAGGAAGGAGAAAGGCTCTAGC 84172
QY 456 SerSerAspGlyGlyLeuPheArgSerArgProAlaHisSerLeuProGlyGlyLeuAsp 475
Db 84173 ATGAAGGACCTGGATGAGTCAAGTGTGGACAGATGGTCCACCCCA-----GAACCT 84226
QY 476 GlyArgValGluProTyr 481
Db 84227 GGCACCTGTCTGACCATTT 84244

Search completed: December 10, 2002, 01:12:13
Job time : 3382.68 secs

GenCore version 5.1.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:57:35 ; Search time 249.44 Seconds
(without alignments)
4550.213 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLQLSQPQSWTDD.....AEFYRLWSVDHGEQSVVTP 504

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141353_2242/app_query.fasta.1.1422
-DB=N-Geneseq.101002 -qfmt=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CNG 1.1.113 @runat_04122002_141353_2242 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -JONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq.101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2580	100.0	1515	21	AAZ48861 Human TAB1 coding
2	2580	100.0	1560	18	AAZ48861 Human TAB1 (TAK1) b
3	2580	100.0	1560	21	AAZ48861 Human TAB1 encodin
4	2580	100.0	1560	21	AAZ48861 Human TAB1-1 nucleo
5	2580	100.0	1569	20	AAZ48861 Human TAB1-FLAG en
6	2577	99.9	1568	20	AAZ48861 Human TAB1 encodin
7	2575	99.8	1560	18	AAZ48861 Human TAB1 (TAK1) b
8	543.5	21.1	16877	22	AAZ48861 Human nervous syst
9	543.5	21.1	16877	22	AAZ48861 Human musculoskele
10	409	15.9	696	21	AAZ48861 Human prostate can
11	365	14.1	211	22	AAZ48861 Human breast cell
12	365	14.1	211	22	AAZ48861 Human foetal liver
13	365	14.1	211	22	AAZ48861 Probe #10699 for g
14	365	14.1	211	22	AAZ48861 Human brain expres
15	365	14.1	211	22	AAZ48861 Human bone marrow
16	365	14.1	211	22	AAZ48861 Probe #10035 for g
17	365	14.1	211	22	AAZ48861 Probe #13988 used
18	365	14.1	211	22	AAZ48861 Probe #5799 used t
19	365	14.1	211	24	ABG13376 Human genome-deriv
20	323.5	12.5	455	22	ABA42092 Human breast cell
21	323.5	12.5	455	22	ABA52514 Human foetal liver
22	323.5	12.5	455	22	ABA22303 Probe #769 for gen
23	323.5	12.5	455	22	AAK00778 Human brain expres
24	323.5	12.5	455	22	AAK26230 Human bone marrow
25	323.5	12.5	455	22	AAI10863 Probe #796 for gen
26	323.5	12.5	455	22	AAI32123 Probe #809 used to
27	323.5	12.5	455	22	AAI00787 Human genome-deriv
28	323.5	12.5	455	22	AAK00818 Human breast cell
29	271.5	10.5	467	22	ABA42093 Human foetal liver
30	271.5	10.5	467	22	ABA52515 Probe #770 for gen
31	271.5	10.5	467	22	ABA22304 Human brain expres
32	271.5	10.5	467	22	AAK00779 Human bone marrow
33	271.5	10.5	467	22	AAK26231 Probe #797 for gen
34	271.5	10.5	467	22	AAI10864 Probe #810 used to
35	271.5	10.5	467	22	AAI32124 Probe #779 used to
36	271.5	10.5	467	22	AAI00788 Human genome-deriv
37	271.5	10.5	467	24	ABG00819 Human breast cell
38	261	10.1	194	22	ABA47247 Human foetal liver
39	261	10.1	194	22	ABA65132 Probe #10700 for g
40	261	10.1	194	22	ABA32234 Human brain expres
41	261	10.1	194	22	AAK13551 Human bone marrow
42	261	10.1	194	22	AAK39290 Probe #10036 for g
43	261	10.1	194	22	AAI20103 Probe #13989 used
44	261	10.1	194	22	AAI45303 Probe #5800 used t
45	261	10.1	194	22	AAI05809

ALIGNMENTS

RESULT 1
AAZ48861
ID AAZ48861 standard; cDNA; 1515 BP.
XX
AC AAZ48861;
XX
DT 24-MAR-2000 (first entry)
XX
DE Human TAB1 coding sequence.

Human; TAB1; XIAP; X-linked inhibitor of apoptosis protein; TGF-beta;
transforming growth factor-beta activated kinase 1; monocytic migration;
TAK1 binding protein 1; extracellular matrix protein production;
cell growth inhibitor; beta-amyloid protein deposition;
immunosuppression; Transforming growth factor-beta; ds.
Homo sapiens.
OS
XX

PN JP1326328-A.
 XX 26-NOV-1999.
 XX 13-MAY-1998; 98JP-0130378.
 XX 13-MAY-1998; 98JP-0130378.
 PA (MATS/) MATSUMOTO K.
 XX MPI: 2000-078337/07.
 DR P-PSDB; AAV59450.
 XX Screening a substance which inhibits combination of the X-linked
 PT inhibitor of apoptosis protein -
 PS Claim 2; Page 25-26; 43pp; Japanese.
 XX
 CC This sequence encodes the human TAB1 protein.
 CC The invention relates to a method for screening a substance inhibiting
 CC the formation of a complex between XIAP and TAB1, in which X-linked
 CC inhibitor of apoptosis protein (XIAP), transforming growth factor-beta
 CC activated kinase 1 (TAK1) binding protein 1 (TAB1) and a substance to be
 CC tested are contacted with each other and then the presence or formation
 CC of a complex between XIAP and TAB1 is detected. The substance can be used
 CC as a drug for extracellular matrix protein production enhancement, cell
 CC growth inhibition, monocyte migration, physiologically active substance
 CC induction, immunosuppression, and beta-amyloid protein deposition. A
 CC substance inhibiting the formation of a complex between TAB1 and XIAP as
 CC well as between XIAP and TGF-beta (Transforming growth factor-beta) type
 CC I and/or type II receptor is useful as a drug.
 XX
 SQ Sequence 1515 BP; 323 A; 457 C; 463 G; 272 T; 0 other;

Alignment Scores:
 Pred. No.: 8, 65e-185 Length: 1515
 Score: 2580.00 Matches: 504
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 21 Gaps: 0

US-09-830-144-4 (1-504) X AAZ48661 (1-1515)

QY 1 MetAlAlaGlnArgSerLeuGlnSerGlnGlnProSerTrpThrAsp 20
 Db 1 ATGGCGGCGAGAGAGAGCTTGCTGCAGAGAGACAGCAAGCTGGACAGATGAC 60
 QY 21 LeuProLeuCyHisLeuSerGlyValGlySerAlaSerAsnArgSerTrpSerIleAsp 40
 Db 61 CTGGCTCTCTGCACTCTCTGGGGTTGGCTCAGCTCCAAACCGCAGCTACTCTGCTGAT 120
 QY 41 GlyLysGlyThrGlnSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
 Db 121 GGCAGAGGCACTAGAGCCACCGCCAGAGAGACAGCTGGCTCAAGTTCAGAGTGAAGC 180
 QY 61 AsnCyPheLeuTrpGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
 Db 181 AACTGCTTCTCTATGGGGCTTCAACGGCTATGATGGCAACCGATGACCAACTTCGTG 240
 QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlnGlnLeuAsnAlaGlnIleAlaGlnVal 100
 Db 241 GCCCAGGCGCTGTCGAGAGCTCTGCTGGGCACTGAAATGCCAGACCGCCGAGGCC 300
 QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGlnArgSerPheLeuGlnSer 120
 Db 301 GATCTCGGGGTGCTGCTGCAGAGCTTGAATGTGTGGAGAGGAGCTTCTCGAGTCC 360
 QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
 Db 361 ATTGACGACCGCTTGGCTGAGAGGCAAGCTTCAGTCCCAATTGGCCAGAGGAGATCCCT 420
 QY 141 GlnHisGlnLeuProProGlnTrpGlnLysIleLeuGlnLysArgLeuLysThrLeuGlnArg 160

Db 421 CAGCACCAAGCTGCTCTCTCAGTATCAGAGATCTTGAAGAGATCTCAAGAGCTTAGAGAG 480
 QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTrpVal 180
 Db 481 GAAATTTTCGGAGAGGGCATGGCCGTTGTGGCGTCTCTCAACAACAAGCTCAAGCTC 540
 QY 181 AlaAsnValGlyThrAsnAlaLeuLeuCyLeuSerThrValAspGlyLeuGlnVal 200
 Db 541 GCCAATGTCGATCAACCCGTGACCTTTATGCAATTCGACATGATGAGGTTCAGAGTG 600
 QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGlyLeuPheArgLeuSerGln 220
 Db 601 ACACAGCTGAACGTGGACCCACACACAGAGAAAGAGATAGCTCTTCCGCTTTGGCAG 660
 QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLysCysGlyGlnGlnSerThr 240
 Db 661 CTGGGCTTGGATGGTGGAAAGATCAAGCAGTGGGGATCATCTGTGGGCGAGAGACACC 720
 QY 241 ArgArgIleGlyAspTrpLysValLysTrpGlyTyrThrAspIleAspLeuLeuSerAla 260
 Db 721 CGGGGATCGGGATTCAAGGTTAAATATGCTTACACGAGCATTTGACTTCACGCT 780
 QY 261 AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyValAlaGlnProLeuAspGly 280
 Db 781 GCCAAGTCCAAACCAATCATCGAGAGCCAGAAATCCATGGGCGACAGCCGCTGGATGG 840
 QY 281 ValThrGlyPheLeuValLeuMetSerGlyLysLeuTrpLysAlaLeuGlnAlaAlaHis 300
 Db 841 GTACGGGCTTCTTGGTGTGATGTGGAGGGTTGTAACAAGGCCCTTAGAGGACGCCAT 900
 QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
 Db 901 GGGCTGGGCAAGGCAACGAGAGATTGCTGCATGATGACATGATTTGGCAAGCAG 960
 QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
 Db 961 ACCTCCTCGAGCGAGAGTGGCCAGAGCCGCTGTGAGACCGGTTGAAGCCATCCACAGCAG 1020
 QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
 Db 1021 ACCTTCGCAAGTGGTGGGAGCGTGCAGAGTTCGCCCCCGGCACGAGCATGACCTTG 1080
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db 1081 CTAGTGAAGAACTTTGGCTACCGCTGGGGAATGAGCCAGCCACACCGAGCCAGGCC 1140
 QY 381 ProAlaIleAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
 Db 1141 CCAGCTGCAAGAGACAGAGTACCTGTGTCTGTGCAATCTCCAGCGCCGACAGAGACC 1200
 QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyMetValAsnGly 420
 Db 1201 AGCAAGACACAGCGTACCTCTCCCTGTGATGCTCCCAAGGCGCAGATGGTCAACGGG 1260
 QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
 Db 1261 GCTCACAGTGTCTTCCCTCGAGAGAGCCACCCCACTTCACCAACCAACCCGAGACC 1320
 QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerAspGlyGly 460
 Db 1321 TTAACTCTGCAATCCACCAACGACACAGACAGACAGAGCTTCAGACTTGCAGAGGC 1380
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
 Db 1381 CTCTTCCGCTCCGGCCCGCCCACTCGCTCCGCGCTGCGAGAGAGCGTGTGTGAGCCC 1440
 QY 481 TyrValAspPheAlaGluLysPheTyrArgLeuTrpSerValAspHisGlyGlnGlnSerVal 500
 Db 1441 TATGTGACTTGTCTGATTTTACCGCTTGAGAGCTGAGCATTGCGAGAGAGAGCGTG 1500
 QY 501 ValThrAlaPro 504


```
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAspArgValLysArgLLeHisSerAsp 340
DB 990 ACCTCCCTGACGACGAGGCGCCGCTGCGAGCCGGGTGAAGGCAATCCACAGCGAC 1049
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGlnuAspMetThrLeu 360
DB 1050 ACCTTGCGCAGTGGTGGGAGCGCTGCGAGTTCTGCCCCGCGACAGACATGACCCCTG 1109
QY 361 LeuValArgAsnPhgGlyTyrProLeuGlyGlyMetSerGlnProThrProSerProAla 380
DB 1110 CTAGTACGAGAACTTGTGCTAACCCGCTGGCGGAATATAGCCAGCCACCGAGCCAGCC 1169
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnInsThr 400
DB 1170 CCGAGCTGACAGAGACGAGTATACCTGTGTCTGTCCATCTCCAGCCGCGCCAGACACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1230 AGCAAGACCCAGCGCTGACCTCTCCCTGTTCATGCCCTCCACAGGCGCAATGGTCMAAG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGlnuAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1290 GCTCACAGTGCCTTCCACCTCGACGAAAGCCACCCCACTCACAACCAAGCCGAGCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
DB 1350 TTAACTCTGACGTCACCAACACGACGACGACGACGACGACGACGACGACGACGACGAC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlyuAspGlyArgValGluPro 480
DB 1410 CTCTTCGCTCCCGGCGCCGCCCTCGCTCCCGCTCGGCGAGAGACGCTGCTTTAGCCCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTyrPheSerValAspHisGlyGlnGlnSerVal 500
DB 1470 TAATGTGACCTTGTGCTAGTTTACCGCTCTGTGAGCGGTGAGACCATGGCAGACGCTG 1529
QY 501 ValThrAlaPro 504
DB 1530 GTGACAGCACCG 1541

RESULT 3
AAK56278
ID AAK56278 standard; DNA; 1560 BP.
XX
AC AAK56278;
XX
DT 21-JUL-1999 (first entry)
XX
DE Human TAB1 encoding DNA.
XX
KW Human; TAB1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 30..1544
FT tag= a
XX
PN WO9921010-A1.
XX
PD 29-APR-1999.
XX
PF 22-OCT-1998; 98WO-JP04796.
XX
PR 22-OCT-1997; 97JP-0290188.
XX
PA (CHUS ) CHUGAI SEIYAKU KK.
XX
PI Ohtomo T, Ono K, Tsuchiya M;
XX
DR WPI: 1999-312645/26.
DR P-PSDB: AAI09541.
```

```
XX
PT Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX
PS Example 1; Page 143-147; 195pp; Japanese.
XX
CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors or activators, or cell proliferation prevention inhibitors or
CC activators, or monocyte migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be
CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes human TAB1.
XX
SQ Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;

Alignment Scores:
Pred. No.: 8.97e-185 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-4 (1-504) x AAK56278 (1-1560)
QY 1 MetaAlaArgArgSerLeuLeuGlnSerGluGlnProSerTPrThrAsp 20
DB 30 ATGGCGGCGAGAGAGAGACTTGTGACAGTGAAGCAGACGACCAAGCTGGAATGAC 89
QY 21 LeuProLeuCySHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 90 CTGCCTCTGTGCAACCTCTCTGGGGTTGGCTCAGCTCCCAACCGGCACTCTGCTGAT 149
QY 41 GlyLysGlyThrGlnSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
DB 150 GGCAGAGGCACTGAGACCAACCCGACAGACGCTGGCTCAAGTTCAGAGATGGAAC 209
QY 61 AsnCySPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPhaVal 80
DB 210 AACTGCTTCTGTAATGGGGCTTTCACAGGCTTAATGAGCAACCGAATGACCACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
DB 270 GCCAGCGGCTGTCCGAGAGACTCTGCTGGCCAGCTGAATCCGAGACCGCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnUser 120
DB 330 GATGTGGCGGTGTGTGCTGACAGGCTTCGATGTGTGTGAGAGGAGCTTCCCTGGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 390 ATTGACACCGCTTGGCTGTGAAGGCAAGCTTCAGTCCGAAATGGCAAGGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
DB 450 CAGACACGCTGCTCCTCGATCAAGAGATCTTGAAGACATCAAGACGTTAGAGAG 509
QY 161 GluLysSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 510 GAAATTTCCGGAGGGGCGCATGGCGCTTGTGGCGGCTTCTCAACCAAGCTCTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
DB 570 GCCAATTCGGTAACAACCGTGCATTTTATGCAAAATGCAAGTGAAGTGGTGGAGTG 629
```



```
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
|||
Db 630 ACACAGCTGAACGTGGACACACACAGAGACGAGATGAGCTCTTCGCTCTTCGCAG 689
QY 221 LeuGlyLeuAspAlaGlyIysIleLysGlnValGlyIleIleCysGlyGlnGluSerThr 240
|||
Db 690 CTGGCTTGGATCTGCTGAAAGATCAAGACAGGTGGGATCATCTGTGGGCGAGAGACCC 749
QY 241 ArgArgIleGlyAspTyrIysValIysTyrGlyTyrThrAspIleAspLeuSerAla 260
|||
Db 750 CGCGCATCGGGATTTACAGGTTAAATATGCTACCGGACATTGACCTTCTCAGCGCT 809
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyValaGlnProLeuAspGly 280
|||
Db 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCATGGGGCCACAGCCGTGGATGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrIysAlaLeuGluAlaHis 300
|||
Db 870 GTACGGGCTTCTTGGTGTGCTGATGTCGAGGGGTTGTACAAGGCCCTAGAGCAGCCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
|||
Db 930 GGGCTTGGGCGAGCCACACAGAGATTGTCGATGATTGACACTGAGTTTGCCRAAGCG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
|||
Db 990 ACCTCCCTGGACGAGTGGGCCAGGCCGTCGTGGACCGGGTGAAGCGCATCCACAGCGAC 1049
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
|||
Db 1050 ACCTTGGCAGTGGTGGGAGCGTGCAGGTTCTGCCCGCGGCAGGACATGACCCCTG 1109
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
|||
Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGCGGAAATGAGCGAGCCACCGAGCCAGCC 1169
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
|||
Db 1170 CCAGCTGCAGGAGCAGAGTGATACCCCTGTCTGTGCCATATCTCCAGCGCCAGACACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
|||
Db 1230 AGCAAGACCAACGCGTGACCTCTCCCTTGTCTATGCCCTCCCGGGCCAGATGGTCAACGGG 1289
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
|||
Db 1290 GCTCACAGTGTCTCCACCTGGACGAGCCACCCCTCCACCAACCAAGCCCGACCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerGlyGly 460
|||
Db 1350 TTAACCTTCAGTCCACCAACACGACACGACGACGAGCAGCTCCAGCTTCGACGGAGGC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGluAspGlyArgValGluPro 480
|||
Db 1410 CTCTTCGCTCCCGGCGCGCCACCTCGCTCCCGCTGGCGAGACGCTGCTGTGAGGCC 1469
QY 481 TyrValAspPheAlaGluPheTyrArgLeuThrSerValAspHisGlyGluGlnSerVal 500
|||
Db 1470 TATGTGGACTTGTGAGTTTATCCGCTCTGGAGCGTGGACCATGGCGAGCAGCGGTG 1529
QY 501 ValThrAlaPro 504
|||
Db 1530 GTACAGCACCG 1541
RESULT 4
AAA39106
ID AAA39106 standard; DNA: 1560 BP.
XX
AC AAA39106;
XX
DT 04-SEP-2000 (first entry)
XX
```

```
DE Human TAB-1 nucleotide sequence SEQ ID NO:3.
XX
KW Human; TAK-1; TAB-1; mitogen activated protein kinase; MAPK;
KW screening; signal transduction; inhibition; inflammatory cytokine;
KW IL-1; interleukin 1; TNF; tumour necrosis factor; inflammation;
KW antiinflammatory; suppression; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 30..1544
FT /*tag= a
FT /product= "TAB-1"
XX
XX WO200023610-A1.
XX
XX 27-APR-2000.
XX
XX 21-OCT-1999; 99WO-JP05817.
XX
XX 21-OCT-1998; 98JP-0299962.
XX
XX (CHUS ) CHUGAI SEIYAKU KK.
XX
XX Tsuchiya M, Ohtomo T, Sugamata Y, Matsumoto K;
XX
XX WPI; 2000-339707/29.
XX
XX P-PSDB; AAY91001.
XX
XX Method for screening inhibitors of TAK1 signal transduction for
XX suppression of inflammatory cytokine production and use as
XX antiinflammatory agents -
XX
XX Disclosure; Page 85-90; 100pp; Japanese.
XX
XX The present invention describes a method for screening compounds for
XX inhibition of inflammatory cytokine signal transduction by contacting
XX the sample with TAK1 and its receptor TAB1 and selecting for inhibition
XX of TAK1/TAB1 binding. Also described is a method for screening compounds
XX for inhibition of inflammatory cytokine signal transduction in which the
XX inhibition of TAK1 phosphorylation is selected for; and drug
XX compositions for the treatment of inflammatory disorders containing as
XX active component an inflammatory cytokine signal transduction inhibitor.
XX TAK1 is an essential component of the signalling process which results
XX in release of inflammatory cytokines such as interleukin-1 (IL-1),
XX IL-10, tumour necrosis factor (TNF) and IL-6. The methods can be used
XX for the selection of effective antiinflammatory agents. The present
XX sequence encodes human TAB-1, which is used in the exemplification of the
XX present invention.
XX
XX Sequence 1560 BP; 332 A; 469 C; 480 G; 279 T; 0 other;
XX
Alignment Scores:
Pred. No.: 8.97e-185 Length: 1560
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-830-144-4 (1-504) x AAA39106 (1-1560)
QY 1 MetaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
|||
Db 30 ATGGCGGCGCAGAGGAGGAGCTTCTGCAGAGTGAAGCAGCCAGCTGGACAGATGAC 89
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
|||
Db 90 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTCCAAACCCAGCTACTCTGCTGAT 149
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
|||
Db 150 GGCAAGGGCACTGAGAGCCACCCGCGAGAGGACAGCTGGCTCAAGTTTCAGGAGTGAGAAC 209
```

```

QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 210 AACTGCTTCTCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGCAACCTTCGGT 269
QY 81 AlaGlyArgLeuSerAlaGlyLeuLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGlnVal 100
Db 270 GCCCAGCGGCTGTCCGCAAGACTCTCTGGGCGCACTTAATGCCAGCAACCCCGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValAlaGlnArgSerPheLeuGlnSer 120
Db 330 GATGTGCGGGGTGTGCTGTGCAAGGCTTCGATGTGTGGAGAGGAGCTTCCTGGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGlnGlyValPro 140
Db 390 ATTGAAGAGCGCTTGGCTAGAGAGCGAACCTCCAGTCCCAATTGCCAAGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArg 160
Db 450 CAGCACCAGCTGCTCTCTCAAGTATCCAGAGATCTTGAAGACTCAAGACGTTAAGAGG 509
QY 161 GlnIleSerGlyValAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 510 GAAATTCGGAGGGGCGCATGGCCGTTGTGGCGTCTTCTCAACAACAAGCTTACGCT 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 570 GCCAATGTGCTGATCAAAACCGTGCACCTTTATGCAATCCACAGTGGATGGCTGCAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGlnAsnGlnAspGlnLeuPheArgLeuSerGln 220
Db 630 ACACACTTAACGCTGACACACACACAGAGACGAGAGTCTCTCCGTCTTCGACG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnLysSerThr 240
Db 690 CTGGGCTTGATGCTGGAAGAATCAAGCAGTGGGATCATCTGTGGGAGAGAGCACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 750 CGCGGATGCGGAGATTACAAGGTAAATATGGCTACACGACATTGACCTTCTCCGCGCT 809
QY 261 AlaLysSerLysProIleIleAlaGlnProGlnIleHisGlyAlaGlnProLeuAspGly 280
Db 810 GCCAAGTCCAAACCAATCTCCGACAGCCAGAAATCCATGGGCGCACGCGCTGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaAlaHis 300
Db 870 GTGACGGGCTTCTGGTGTGATGTCGAGGGGTTGTACAAGCCCTGAGAGCAGCCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnLysIleAlaAlaMetIleAspThrGlnPheAlaLysGln 320
Db 930 GGGCGCTGGGCAAGCCAAACGAGAGATGTCTGCATGATTCACATGAGTTGGCCAAAGAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
Db 990 AACTCTCTTGAAGCAGTGGCCCAAGCGCTGTGACCGGGTAAACGATCCACAGCCAGAC 1049
QY 341 ThrPheAlaSerGlyGlyGlnArgAlaArgPheCysProArgHisGlnAspMetThrLeu 360
Db 1050 ACCTTCGCGCAGTGTGGGAGCGTGCAGGTTCTGCCCCCGGACAGAGACATGACCTTG 1109
QY 361 LeuValArgAsnArgGlyTyrProLeuGlyGlnMetSerGlnProThrProSerProAla 380
Db 1110 CTAAGTGAAGAACTTTGGCTACCCGCTGGGCGAAATGAGCCACCCACGCGCCACACC 1169
QY 381 ProAlaAlaGlyArgValTyrProValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
Db 1170 CCAAGCTCAGAGGAGAGAGACTGTACCTGTGTGTGTCATACTCCAGGCGCCAGAGACC 1229
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
Db 1230 AGCAAGACAGCGTGAACCTCTCCTTGTTCATGACCTCCACAGGCGCAATGTGTCAACGGG 1289

```

```

QY 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrIleuThrAsnGlnSerProThr 440
Db 1290 GCTCACAGTGTCTTCACCTCGGAGCAAGCCACCCCACTTCACCAACCAACCCGAGCC 1349
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
Db 1350 TTAACTTCGAGTCCACCAACGACGACGACGAGAGAGAGCTTCAGCTTGAAGAGCC 1409
QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGlnPro 480
Db 1410 CTTTCGCTCCCGGCGCCGCACTGCTCCGCTCGGAGGAGACGCTGTGTAGGCC 1469
QY 481 TyrValAspPheAlaGlnPheTyrArgLeuTyrPheValAspHisGlyGlnGlnSerVal 500
Db 1470 TATGTGACTTGTGCTGATTCACGCTTCGAGCGTGAACCATGCGAGCAGACGCTG 1529
QY 501 ValThrAlaPro 504
Db 1530 GTGACAGCACCG 1541

RESULT 5
ID AAX56282
XX AAX56282 standard; DNA; 1569 BP.
AC AAX56282;
XX 21-JUL-1999 (first entry)
DT 21-JUL-1999 (first entry)
DE Human TAB1-FLAG encoding DNA.
XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX Homo sapiens.
OS Synthetic.
OS Key Location/Qualifiers
FH CDS 7..1560
FT /*tag= a

MO9921010-At.
PD 29-APR-1999.
XX 22-OCT-1998; 98WO-JP04796.
XX 22-OCT-1997; 97JP-0290188.
XX (CHUS ) CHUGAI SEIYAKU KK.
XX Ohtomo T, Ono K, Tsuchiya M;
XX WPI; 1999-312645/26.
XX DR P-PDSB; AAY09546.
XX PT Screening for TGF- beta inhibitory substances, which are useful as
XX drugs for treatment of diseases relating to its disorder
XX Example 1; Page 159-163; 1995pp; Japanese.

CC A method has been developed for screening for substances which inhibit
CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
CC comprises: (a) contacting the polypeptide in the presence of a sample;
CC and (b) detecting the amount of bound polypeptide, in which the sample
CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
CC growth factor (TGF)-beta inhibitory substances can be used in drugs for
CC indications e.g. as TGF-beta signal transduction inhibitors or
CC activators, or extracellular matrix protein production enhancement
CC inhibitors, or monocytic migration inhibitors or activators, or
CC physiological activity induction inhibitors or activators, or
CC immunosuppression inhibitors or activators, or amyloid beta protein
CC precipitation inhibitors or activators, and such substances can also be

```

CC inhibitors of the TAK1 polypeptide function, particularly kinase
CC activity. The present sequence encodes TAB1-FLAG from an example of
CC the present invention.

XX Sequence 1569 BP; 343 A; 466 C; 476 G; 284 T; 0 other;

Alignment Scores:

Pred. No.: 9,04e-185 Length: 1569
Score: 2580.00 Matches: 504
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-09-830-144-4 (1-504) x AAX56282 (1-1569)

QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerThrThrAspAsp 20
DB 7 ATGGCGGCGCAGAGGAGGCTTCTGCAGAGTGAGCAGCAGCAAGCTGGACAGATGAC 66
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 67 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCCTCCAAACCGCAGCTACTCTGTCTGAT 126
QY 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
DB 127 GGCAAGGGCACTGAGAGCCACCCGCCAGAGACAGCTGGCTCAAGTTTCAGGAGTGAGAAC 186
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 187 AACTGCCTTCCTGTATGGGCTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 246
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
DB 247 GCCCAGCGCTGTCGAGAGCTCTCTGCGCCAGCTGAATGCCGAGCAGCCGAGGCC 306
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
DB 307 GATGTGGCGGTGTGCTGCTGCAGGCTTCGATGTGTGTGGAGGAGCTTCCTGGAGTCC 366
QY 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 367 ATTGACACGCTTGGCTGAGAGGCAAGCCTCCAGTCGCAATGGCAGAGGAGTCCCT 426
QY 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
DB 427 CAGCACCGCTGCTCTCTAGTATCAGAGATCTCTGAGAGACTCAAGAGCTTAGAGAGG 486
QY 161 GluIleSerGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
DB 487 GAAATTTCCGGAGGGGCATGGCGTTGTGGCGTCTCTTCAACAACAAGCTCTACGTC 546
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
DB 547 GCCAATGTCGGTCAAAACCGTGACCTTTTATGCAAAATCGACAGTGGATGGTTGCAGGTG 606
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
DB 607 ACACAGCTGAACGTGGACCAACACACAGAGAGATGAGCTCTTCGTCTTTTCGAG 666
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleCysGlyGlnGluSerThr 240
DB 667 CTGGCTTGGATGCTGGAAGATCAACGAGGTGGGATCATCTGTGGCAGGAGACACC 726
QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
DB 727 CGCGGATCGGGATTACAAGGTAAATATGGCTACACGACATTTGACCTTCTCAGCGCT 786
QY 261 AlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
DB 787 GCCAAGTCCAACAACATCATCTGAGAGGCCAGAAATCCATGGGGCAGCGCTGGATGGG 846
QY 281 ValThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaHis 300

DB 847 GTGACGGGCTTCTGTGCTGATGTCGAGGGGTGTATCAAGGCCCTAGAGCGACCCCAT 906
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaMetIleAspThrGluPheAlaLysGln 320
DB 907 GGGCTGGGCGAGGCCAACAGGAGATTGCTGCGATTGACACTGAGTTTGTCCCAAGCAG 966
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
DB 967 ACCTCTCTGGACCGAGTGGCCAGGCCGCTGTGGACCGGGTGAAGCGCATCCACAGCGAC 1026
QY 341 ThrPheAlaSerGlyGlyGluArgAlaArgPheCysProArgHisGluAspMetThrLeu 360
DB 1027 ACCTTCCGACAGTGTGGGAGCGTGCAGGTTCTGCCCCCGCAGCAGACATGACCCCTG 1086
QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGluMetSerGlnProThrProSerProAla 380
DB 1087 CTAGTGAGGAACTTTGGCTACCCGCTGGGCGAAATGAGCCAGCCACACCGAGCCAGCC 1146
QY 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
DB 1147 CCAGCTGCAGGAGGACGAGTGATCCCTGTGTCTGTGCCATTACTCCAGCGCCAGAGCACC 1206
QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
DB 1207 AGCAAGACCCAGCGTGACCTCTCCCTTGTCTGTCATGCCCTCCAGGGCCAGATGTTCAACGGG 1266
QY 421 AlaHisSerAlaSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThr 440
DB 1267 GCTCAGTGTCTCCACCTGGAGCAGCCACCCCTCAGCTCAGCAACAAAGCCCGACC 1326
QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerSerSerSer 460
DB 1327 TTAACCTTGCACTCCACCAACACGACACGACAGACAGCTCCAGCTCTGACGAGGC 1386
QY 461 LeuPheArgSerArgProAlaHisSerLeuProGlyGluAspGlyArgValGluPro 480
DB 1387 CTCCTCCGCTCCCGGCCCGCCCACTCGCTCCCGCTGGCAGGAGCGTCTGTTGAGCCC 1446
QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGluGlnSerVal 500
DB 1447 TATGTGGACTTTGTGTAGTTTTTACCGCTCTGGAGCGTGGACCATGGCGAGAGCGGTG 1506
QY 501 ValThrAlaPro 504
DB 1507 GTGACAGCACCG 1518
RESULT 6
AAX56310
ID AAX56310 standard; DNA; 1568 BP.
XX AAX56310;
AC AAX56310;
XX 21-JUL-1999 (first entry)
DE Human TAB1 encoding DNA SEQ ID NO:42.
XX Human; TAB1; TAK1; screening; inhibition; TGF-beta;
KW transforming growth factor beta; ss.
XX Homo sapiens.
FH Key Location/Qualifiers
FT CDS 11..1552
FT /*tag= a
PN WO9921010-A1.
XX 29-APR-1999.
XX 22-OCT-1998; 98WO-JP04796.
XX 22-OCT-1997; 97JP-0290188.

XX (CHUS) CHUGAI SEIYAKU KK.
 PA
 XX Ohtomo T, Ono K, Tsuchiya M;
 PI
 XX WPI, 1999-312645/26.
 DR P-PSDB; AA09550.

XX Screening for TGF- β inhibitory substances, which are useful as
 PT drugs for treatment of diseases relating to its disorder
 XX

PS Example 13; Page 182-186; 195pp; Japanese.

XX A method has been developed for screening for substances which inhibit
 CC the binding of TAK1 polypeptide to TAB1 polypeptide. The method
 CC comprises: (a) contacting the polypeptide in the presence of a sample;
 CC and (b) detecting the amount of bound polypeptide, in which the sample
 CC can be pre-mixed with TAK1 or TAB1 polypeptide first. The transforming
 CC growth factor (TGF)- β inhibitory substances can be used in drugs for
 CC indications e.g. as TGF- β signal transduction inhibitors or
 CC activators, or extracellular matrix protein production enhancement
 CC inhibitors, or monocytic migration inhibitors or activators, or
 CC activators, or physiological induction inhibitors or activators, or
 CC immunosuppression inhibitors or activators, or amyloid β protein
 CC precipitation inhibitors or activators, and such substances can also be
 CC inhibitors of the TAK1 polypeptide function, particularly kinase
 CC activity. The present sequence encodes human TAB1.

XX Sequence 1568 BP; 339 A; 472 C; 477 G; 280 T; 0 other;

Alignment Scores:

Pred. No.:	1,52e-184	Length:	1568
Score:	2577.00	Matches:	503
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	99.80%	Mismatches:	0
Query Match:	99.88%	Indels:	0
DB:	20	Gaps:	0

US-09-830-144-4 (1-504) X AAX56310 (1-1568)

QY 1 MetAlaIaGlnArgArgSerLeuGlnSerGluGlnProSerTrpHisAsp 20
 DB 38 ATGGCGCGGAGAGGAGGAGCTTGCGAGAGGAGCAGAGCAAGCTGGACAGATGAC 97
 QY 21 LeuProLeuCyHisIleuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaasp 40
 DB 98 CTCGCTCTCTGCCACCTCTCTGGGTTGGTCAAGCTCCACCGCAGTACTGCTGAT 157
 QY 41 GlyLysGlyThrGlnSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
 DB 158 GGCAGAGGCGACTGAGAGCCACCGCAGAGCAGCTGGTCAAGTTCAGGAGTGAGAAC 217
 QY 61 AsnCySPheLeuArgValPheAsnGlyTyrAspGlyAlaAsnArgValThrAsnPheVal 80
 DB 218 AACTGCTCTCTGATGGGGCTTCAACGGCTATGATGGCAACGAGTACCACTTCGTG 277
 QY 81 AlaGlnArgLeuSerAlaGlnLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnAla 100
 DB 278 GCCCAGGCGGTCTCCGAGAGCTCTGCTGGGCGACCTGATGCCAGACCGCGAGGCC 337
 QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGlnArgSerPheLeuGlnSer 120
 DB 338 GATGTGGGGGTGGCTGCGAGGCTTCGATGTGTGAGAGGAGGAGCTTCTCGAGTCC 397
 QY 121 ILAspAspAlaLeuAlaGlnLysAlaSerLeuGlnSerGlnLeuProGluValPro 140
 DB 398 ATTGACGACCGCTTGGCTGAGAGGCAAGCTCCAGTCGAAATTGCGAGGGAGTCCCT 457
 QY 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArg 160
 DB 458 CAGCACGAGCTGCTCTCTCAGTATCAGAGATCTCTTGAGAGACTCAAGACGTTAGAGAG 517

QY 161 GlnIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
 DB 518 GAATTTCCGGAGAGGCGCAGGCGGTTGTGGCGGCTCTTCAACAACAGCTTACGTC 577
 QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCyLysSerThrValAspGlyLeuGlnVal 200
 DB 578 GCCAATTCGGTACCAACCGTGCACCTTTATATGCAATTCGACAGTGTAGGTTCAGGTC 637
 QY 201 ThrGlnLeuAsnValAspHisThrThrGlnSerGlnAspGlnLeuPheArgLeuSerGln 220
 DB 638 ACACAGCTGAACGTGGCCACACACAGAGAACAGAGATGAGCTCTTCGCTTCCTGCGAG 697
 QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCyGlnGlnIleSerThr 240
 DB 698 CTGGGCTTGAATGCTGAAAGATCAAGACAGTGGGATCATCTGTGGGAGAGAGACACC 757
 QY 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAla 260
 DB 758 CGGCGATCCGGGATTACAGAGTTAAATATGCTACACCGACATTGACCTTCACGGCT 817
 QY 261 AlaLysSerLysProIleAlaGluProGlnIleHisGlyAlaGlnProLeuAspGly 280
 DB 818 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGCGACAGCGCTGATGGG 877
 QY 281 ValThrGlyPheLeuValIleuMetSerGlnGlyLeuTyrLysAlaLeuGlnAlaHis 300
 DB 878 GTGACGGGCTTCTTGCTGATGTGAGAGGGTTGACAAAGCCCTTAGAGGAGCCCAT 937
 QY 301 GlyProGlyGlnAlaAsnGlnIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
 DB 938 GGGCTGGCAGGCGCAACCAAGAGATTGCTGCATGATGATGACCTGATGGTTCGCAAGCAG 997
 QY 321 ThrSerLeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAsp 340
 DB 998 ACCCTCCCTGGACCGCAGTGGCCAGGCGCTGTGATGAGCGGTGAAGGCATCCACACGAC 1057
 QY 341 ThrPheLeuSerGlyGlyValArgAlaArgPheCysProArgHisGlnAspMetTrpLeu 360
 DB 1058 ACCTTCGCGAGTGTGGGAGCGTGTCCAGGTCTTGCCCGGCGACAGAGACATGACCTTG 1117
 QY 361 LeuValArgAsnPheGlyTyrProLeuGlyGlnMetSerGlnProThrProSerProAla 380
 DB 1118 CTATGAGGAACCTTTGGCTACCGGCTGGCCAAATGAGCCACACCGACCCAGGCC 1177
 QY 381 ProAlaAlaGlyArgValTyrProValSerValProTyrSerSerAlaGlnSerThr 400
 DB 1178 CCAGCTCGAGGAGGACGAGTACCTCTGTCTGTGCTCATATCCAGCGCCAGAGCACCC 1237
 QY 401 SerLysThrSerValThrLeuSerLeuValMetProSerGlnGlyGlnMetValAsnGly 420
 DB 1238 AGCAAGACCAAGCGTACCTCTCTCTGTGATGCTCCAGAGGCCAGATGATCAACGGG 1297
 QY 421 AlaHisSerAlaSerThrLeuAspGlnAlaThrProThrLeuThrAsnGlnSerProThr 440
 DB 1298 GCTCAGAGTGTCTCAACCTGAGAGAACCAACCCACCTTCACCAACCAACCCGAGCC 1357
 QY 441 LeuThrLeuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerAspGlyGly 460
 DB 1358 TTAACCTCGCAGTCCACCAACACGACACAGCAGCGAGCTCCAGCTTCGACGGAGGC 1417
 QY 461 LeuPheArgSerArgProAlaHisSerLeuProProGlyGlnAspGlyArgValGluPro 480
 DB 1418 CTCTTCGCTCCGCGCCCGCCACCTCTCCCGCTGGCGAGAGCGTCTGTGAGCCC 1477
 QY 481 TyrValAspPheAlaGluPheTyrArgLeuTrpSerValAspHisGlyGlnGlnSerVal 500
 DB 1478 TATGTGACCTTGTGATGATTTTACCGCTCTGAGCGTGAAGCTGACGAGAGAGGCTG 1537
 QY 501 ValThrAlaPro 504
 DB 1538 GTGACAGCACCG 1549

RESULT 7

AAT91178
ID AAT91178 standard; cDNA; 1560 BP.
XX AC AAT91178;
XX 14-APR-1998 (first entry)
XX Human TAB1 (TAK1 binding protein) cDNA.
XX TAB1; TAK1 binding protein; transforming growth factor-beta;
KW signal transduction; human; ds.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
ID CDS 30..1544
XX FT /*tag= a
XX FT variation 185
XX FT /*tag= b
XX FT /note= "another clone has cytosine at position 185,
FT with codon AGA (Arg) altered to AGC (Ser)"
XX
PN EP803571-A2.
XX
PD 29-OCT-1997.
XX
PF 24-APR-1997; 97EP-0302808.
XX
PR 20-NOV-1996; 96US-0752891.
PR 24-APR-1996; 96JP-0126282.
PR 28-OCT-1996; 96JP-0300856.
XX
PA (UENO/) UENO N.
XX
XX Matsumoto K, Nishida E;
XX WPI; 1997-515318/48.
DR P-PSDB; AAW26707.
XX
XX DNA encoding TAK1 binding protein TAB1 - member of transforming
PT growth factor beta receptor signal production pathway, which
PT activates TAK-1 kinase activity upon binding
XX
PS Example 5; Page 19-21; 30pp; English.
XX
XX This cDNA clone codes for human TAB1 (see AAW26707), a novel member
CC of the transforming growth factor-beta receptor signal transduction
CC pathway, which activates TAK-1 kinase activity upon binding. To
CC obtain the full-length TAB1 coding sequence, a human kidney cDNA
CC library was screened using as a probe a partial TAB1 cDNA obtained
CC from a yeast two-hybrid assay for proteins that interacted with
CC TAK1. The 5' terminus was identified by 5'RACE. 2 Different
CC clones were sequenced, with cytosine (see AAT91178) and adenine as
CC the 185th nucleotide, respectively, and deposited as FERM BP-5599
CC and FERM BP-5508, respectively. Also claimed are: isolated DNA
CC encoding a protein modified by a substitution, deletion and/or
CC addition of 1 or more amino acids of the 504-residue TAB1 sequence;
CC (2) DNA which can hybridise with the 1560 bp nucleic acid sequence;
CC (3) isolated DNA encoding a protein comprising amino acids 21-579
CC or 437-504 of the 504 TAB1 sequence; (4) DNA encoding a fusion
CC protein comprising an above protein or polypeptide; (5) expression
CC vector comprising an above DNA; and (6) host cell, preferably a
CC mammalian or yeast cell, transformed by the expression vector.
CC Cells expressing TAB1 and TAK1 can be used to screen for TGF-beta
CC signalling pathway inhibitors by contacting the cells with a test
CC compound, and measuring the TAK1 kinase activity.
XX
SQ Sequence 1560 BP; 333 A; 468 C; 480 G; 279 T; 0 other;
Alignment Scores:
Pred. No.: 2,13e-184 Length: 1560
Score: 2575.00 Matches: 503
Percent Similarity: 99.80% Conservative: 0

Best Local Similarity: 99.80% Mismatches: 1
Query Match: 99.81% Indels: 0
DB: 18 Gaps: 0
US-09-830-144-4 (1-504) x AAT91178 (1-1560)
QY 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
DB 30 ATGGCGCGCAGAGGAGGAGCTTGTGCAGAGTGAGCAGCAGCAAGCTGCAGAGTGAC 89
QY 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
DB 90 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTCCAAACCCGAGCTACTCTGCTGAT 149
QY 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
DB 150 GCGAAGGCGCATGAGAGCACCCGCCAGAGGACAGATGGCTCAAGTTTCAGAGGTGAGAAC 209
QY 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
DB 210 AACTGCTTCTGTATGGGGTCTTCAACGGCTATGATGCAACCCGAGTGACCACTTCGTG 269
QY 81 AlaGlnArgLeuSerAlaGluLeuLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
DB 270 GCCCAGCGGCTGTCCGACAGAGCTCTCTGCGGCCAGCTGAATGCCGAGCAGCGGAGGCC 329
QY 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
DB 330 GATGTGCGGCGTGTGCTGCTGCAGGCCCTTTCGATGTGTGGTGAGAGGAGCTTCTCGAGTCC 389
QY 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
DB 390 ATTGACGACGCTTGGCTGACAGGCAAGCCTCCAGTCGCAATTCGCCAGGAGGAGTCCCT 449
QY 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
DB 450 CAGCACCCAGCTGCCTCTCTCAGTATCAGAAGATCCTTCGAGAGACTCAAGACGCTTAGAGAGG 509
QY 161 GluIleSerGlyGlyAlaMetAlaValAlaValAlaValLeuLeuAsnLysLeuTyrVal 180
DB 510 GAAATTTCCGGAGGGGCGCATGGCCGTTGTGGCGGTCTCTTCAACAACAAGCTCTACGTC 569
QY 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
DB 570 GCCAATGTCGGTACAAACCGTGCACCTTTTATGCAAAATCGACAGTGGATGGTTCAGGTG 629
QY 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
DB 630 ACACAGCTGAACGTGGACCAACACACAGAGAACGAGGATGAGCTCTTCCGCTTTTCGAG 689
QY 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIlelleCysGlyGlnGluSerThr 240
DB 690 CTGGCTTGGATGCTGGAAAGATCAAGCAGGTGGGGATCATCTGTGGCGAGGAGAGCACC 749
QY 241 ArgArgIleGlyAspTyrLysValLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
DB 750 CGCGGATCGGGATTAACAAGTTAAATATGGCTACACGACATGACCTTCTCAGCGCT 809
QY 261 AlaLysSerLysProIlelleAlaGluProGluIleHisGlyAlaGlnProLeuAspGly 280
DB 810 GCCAAGTCCAAACCAATCATCGCAGAGCCAGAAATCCATGGGGCAGACGCGCTGGATGGG 869
QY 281 ValThrGlyPheLeuValLeuMetSerGluLysLeuTyrLysAlaLeuGluAlaAlaHis 300
DB 870 GTGACGGCTTCTTTGGTGTGATCTCGAGGGGTGTACAAGGCCCTTAGAGGAGCGCCAT 929
QY 301 GlyProGlyGlnAlaAsnGlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGln 320
DB 930 GGGCTGGGCGAGGCCAACCCAGGAGATTGCTGCGATGATTTGACACTGAGTTGCCAAGCAG 989
QY 321 ThrSerLeuAspAlaValAlaGlnAlaValValAspArgValLysArgIleHisSerAsp 340
DB 990 ACCTCCCTGGACGAGTGGGCCAGGCGCTCTGTGACCCGGGTGAAGCGCATCCACAGCGAC 1049

Qy 341 ThrPheAlaSerGIyGlyGluArgAlaArgPheCysProAlaRHisGluAspMetThrIleu 360
 Db 1050 ACCTTCGCAGTGTGGGGAGCGGTCCAGGTTCTGCCCCCGGACAGACATGACCTCG 1109
 Qy 361 LeuValArgAspPheGlyIyrProLeuGlyIuMetSerGlnProThrProSerProAla 380
 Db 1110 CTAGTGAGGAACCTTGGCTACCCGCTGGGCGAAATGAGCAGCCACACCGAGCCAGCC 1169
 Qy 381 ProAlaAlaGlyGlyArgValTyrProValSerValProTyrSerSerIaGlnSerThr 400
 Db 1170 CCAAGTCGAGGAGGAGCAGAGTGAACCTGTGTCTGTGCTCATCTCCAGCGCCAGAGCACC 1229
 Qy 401 SerIySerValThrIleuSerIleuValMetProSerGlnGlyGlnMetValSengly 420
 Db 1230 AGCAAGACAGCGGTGACCTCTCCCTTGTCAATGCCCCCTCCAGGGCCAGTGTCAACGG 1289
 Qy 421 AlaHisSerAlaSerThrIleuAspGluAlaThrProThrIleuThrAsnGlnSerProThr 440
 Db 1290 GCTCACAGTGTCTCAACCTTGAGAGAGCCACCCCACTTCAACCAAGCCCGAC 1349
 Qy 441 LeuThrIleuGlnSerThrAsnThrHisThrGlnSerSerSerSerSerSerIaGlyIy 460
 Db 1350 TTAACTTCGAGTCACCAACAGCAGCAGCAGCAGCAGCAGCTCCAGCTGACGAGGC 1409
 Qy 461 LeuPheArgSerArgProAlaHisSerIleuProProGlyGluAspGlyArgValGluPro 480
 Db 1410 CTCTTCGCGCTCCGCGCCGCTCGCTCCGCTCGGAGACGCTGTGTGAGCC 1469
 Qy 481 TyrValAspPheAlaGluPheTyrArgLeuTyrSerValAspHisGlyGlnGlnSerVal 500
 Db 1470 TATGTGACTTGTCTGATTTTACCGCTCTGAGCGTGACATGCGAGCAGAGCGTG 1529
 Qy 501 ValThrAlaPro 504
 Db 1530 GTGACGACCG 1541
 RESULT 8
 ABA20494
 ID ABA20494 standard; DNA; 16877 BP.
 XX
 AC ABA20494;
 XX
 DT 23-JAN-2002 (first entry)
 XX
 DE Human nervous system related polynucleotide SEQ ID NO 12825.
 XX
 KW Human; nootropic; neuroprotective; cytosratic; dermatological; viroicide;
 KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
 KW antiparkinsonian; antistickling; antianaemic; antiarthritic; cancer;
 KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
 KW antiallergic; antidiabetic; antilucer; anticonvulsant; antifungal;
 KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200159063-A2.
 XX
 PD 16-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01334.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0186874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.

PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228824.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 08-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0236327.
 PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.

PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 20-OCT-2000; 2000US-0242221.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246509.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0251030.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 03-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-541565/60.
XX

XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,
PT useful for preventing, diagnosing and/or treating nervous system
PT cancers and metastases -
XX

XX Disclosure; SEQ ID NO 12825; 1701pp + Sequence Listing; English.
XX

XX The invention relates to novel genes (ABA11004-ABA21534) and proteins
CC (AB14678-AB18001) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone

CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;

Alignment Scores:

Pred. No.: 2,27e-30 Length: 16877
Score: 543.50 Matches: 295
Percent Similarity: 18.86% Conservative: 43
Best Local Similarity: 16.46% Mismatches: 107
Query Match: 21.07% Indels: 1348
DB: 22 Gaps: 19

US-09-830-144-4 (1-504) x ABA20494 (1-16877)

QY 12 GluGlnGlnProSerTyrThrAspAspLeuProLeuCysHisLeuSerGlyValGlySer 31
DB 1 GAGCAGCAGCCAAAGCTGGACAGATGACCTCTCTCTGCCACCTCTCTGGGGTTGGCTCA 60
QY 32 AlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisProGluAsp 51
DB 61 GCCTCAACCGCAGCTACTCTGCTGATGGCAAGGCACTGAGAGCCACCCGCCAGGAC 120
QY 52 SerTrpLeuLysPhe----- 56
DB 121 AGCTGGCTCAAGTTCAGGTGTGTGCCAGCATTTCTGTGTGGCCCGGGAGTTGTT 180
QY 56 ----- 56
DB 181 GGTTCGCAAGCAAGAAAGACACCCACCTTCGAGCTTTCTCGTATGGCTTCCAGTGAC 240
QY 56 ----- 56
DB 241 ATGTGGCCCGTGAGAGTGGGCTCTGTGCTGTCTTGGCAAGGGCCCTGCTGTATGGGGT 300
QY 56 ----- 56
DB 301 AGCGTGAGCATGGGAGGAGGGTATCCAGAAATGTCATAGCCAGAGTGAATGATGGCTA 360
QY 56 ----- 56
DB 361 AAGCAGGGGACCCAGAGAGGGCCCTGAAAGCTGCAGCTGTGCTGCTTAGTCTCCCCA 420
QY 56 ----- 56
DB 421 ATTCCTTTCCCTTTCTCCCTCCACCTCCGTCGAGACCCCTGGTCTCAGGCCTCCCTGTGCC 480
QY 57 -----ArgSerGluAsnAsnCysPheLeuTyrGlyValPheAsnGlyTyrAsp 72
DB 481 TCTCCCTCTTCAGAGGTGAGAACAACTGCTTCCTGTATGGGGCTTTCACCGGTATGAT 540
QY 73 GlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLeuLeuGlyGln 92
DB 541 GGCAACCGAGTGACCAACTTGTGGCCCGCGGCTGTCCGAGAGCTCCTGTGGGCGAG 600
QY 93 LeuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuLeuGln----- 108
DB 601 CTGAATGCCGACGACGCGAGGCCGATGTGCGGCGTGTGCTGTGCTGCAGGTAATGTCGCCG 660
QY 108 ----- 108
DB 661 GGGCCACAGTGACCCAGCCACATCATGTCCCCCACCACCCAAAGGCTTGGGCCCTGCACCTC 720
QY 108 ----- 108

Db 721 TAGCATGTTGCCAGGGTGTGTGAAGATCTGCGGCCCCCTTCAACCCAGTAGAGAGC 780
Qy 108 ----- 108
Db 781 AGCTCCACGATAGAGCCCCCCCACCAACAGAGATCCAGGACCAGCAGCAGCAGG 840
Qy 108 ----- 108
Db 841 CCTTCATGAGCTGGCAGAGAGGTAAGTCTGTGAGACGCGGGGATTTAGGATGGAGCTT 900
Qy 108 ----- 108
Db 901 GGAGAGAGTGTAGAGTGGGAGCAGGGCAAGGCTGTGTAGAAATGGGTCATTTAGAGT 960
Qy 108 ----- 108
Db 961 ACCCTTTCTTCTCTATGTGTGTCAGGTGCTCAGCTCAGGTGACAGAGCATCCCTGGG 1020
Qy 108 ----- 108
Db 1021 CGTCCAGAGAGACCTTGCCCTCCCTTCTGAGGGGCCGCCCTCATTTGACTGTTC 1080
Qy 108 ----- 108
Db 1081 ACCTTTCTGTGCGCTGTGGCAGTTTCTCTGTGCACTGTCTTTCAGAGCATGTCT 1140
Qy 108 ----- 108
Db 1141 CAGGCCATTTGAGATGAGAGATGGGCTTCTGTTCGCGAGAGGGGGTGTGCAGCCTT 1200
Qy 108 ----- 108
Db 1201 TTCTCCCTTACAGACCTCAGGCTTCATGCGCATGATTCAGCAGATCTCACAGAG 1260
Qy 108 ----- 108
Db 1261 GGGAGAGGTGTGATGTCAACAGTGTCTGGGGCTGTGTGGGTTTGACAGAAAGCTCCC 1320
Qy 108 ----- 108
Db 1321 AGAAGAAGTTGATGATATCATACGAGTGTCACTAGCCTGGCAGTTTCTGAGAGAGTT 1380
Qy 108 ----- 108
Db 1381 GCCTTGAGTGCAGGTTGACTGTCTTACTTACCTAACCCCTGGAACACACTTAACCTCC 1440
Qy 108 ----- 108
Db 1441 GTTGTGTGAGACTGAGGGGGCCAGAGGTCAACACAGCTGGGCTTACGCCAACCTTTGCT 1500
Qy 108 ----- 108
Db 1501 CTTAGCTTCGACGTGTCTCTGCTCCACGCGCTGTGACGTGTGCTCTGTGACAGGTTGC 1560
Qy 108 ----- 108
Db 1561 ACACCCCTTCGAGGCTGAAAGGCTTTGTCAAAACATTGATCTGACGAGAGCGCGGTG 1620
Qy 108 ----- 108
Db 1621 CCTTGACGTGTGGGCCAGAGCGACGACTGACATGTGAAAGCTCCATACACAAAGAAC 1680
Qy 108 ----- 108
Db 1681 TGCAGTGAAGACGCAAGGTGCTGCTGATTAATAGAGCAATTTTGGACCAAGTAC 1740
Qy 109 ----- 113
Db 1741 AGTGTGTTTGAACAGGCTTTGCCCTGTCTGTGTCCCTTCCCTTCCATGTGTGTG 1800
Qy 114 GUAUsetPhenugJuserJleaspaapAlaleuAlglulylalaserJuser 133
|||||

Db 1801 GAGAGAGCTTCCGTGAGTCCATTGAAGAGCCTTGGCTGAGAAAGCAAGCCTCAGTGC 1860
Qy 134 GlnLeuProglu----- 137
Db 1861 CAATTGCCAGA-GGTATTTCCCGAGCCGACACCGAGGGAGTCAAGTCCAGGCCAGCT 1919
Qy 137 ----- 137
Db 1920 TTGCAGAGCATGACTCATCTTCTTTGACATTACTGGGCCAGAGCACAGGCCGTT 1979
Qy 137 ----- 137
Db 1980 AGGAGCAGTTCCGATGGGTGACACTGTGTGTGGCCACAGGTGAGGGACCTCGCTGCT 2039
Qy 137 ----- 137
Db 2040 CTGCTGTGGCAGCTGCATGCTAAAGCTCAGAGTGTGTGGCAGTGTGACATGAGGA 2099
Qy 137 ----- 137
Db 2100 AGGCTTTTCCAGTCTTGTCTGTGAATGGTCCCGCCCTTGTCAACCACTGCTCAC 2159
Qy 137 ----- 137
Db 2160 ATGCGCTACTGCCACCTGACGCGCTTACCTGTGCTGTGATGTGGTCACTCCAGCTT 2219
Qy 137 ----- 137
Db 2220 TCTGGAAGAACCTTAGCCTGAGCTGTAGCGAAGCTTCATCTTGCTGTGCCATGAGAC 2279
Qy 137 ----- 137
Db 2280 ATCTGTGTGTGTCCCCCGGCCCTGTGGATCTTTGCTTTAAGACAAAGGAGGCCG 2339
Qy 137 ----- 137
Db 2340 GCATGTGGCTCAGCGCTGTATATCCAGACACTTGGGAGGCGAGGCGGCGGATCACAA 2399
Qy 137 ----- 137
Db 2400 GGTGAGAGATCAAGACCATCTGTGCTAACATGTGTAACCCGCTCTTAAATAATAC 2459
Qy 137 ----- 137
Db 2460 AAAAAATTAGTGGGCTGTGTGTGTGGCGCTTGTATGTCCAGCTACTCTCGAGGCTGAA 2519
Qy 137 ----- 137
Db 2520 GCAGAGATGTGTGAACCTGGGAGGCGGAGCTTGTATGAGCCGAGATCATGCCACTG 2579
Qy 137 ----- 137
Db 2580 CACTGTAGCCTGGGGGACACAGGAGATTCATCTCAAAAAAAAAAAGTAAGG 2639
Qy 137 ----- 137
Db 2640 AGTGAAGGCGTAGAAGATGTCCAGAGTCTGTCTGCGCTGACCTCTGTATG 2699
Qy 138 -----GlyValProglNhsGlnLeuProglNtyrGlnLysIleLeuGluArGlu 155
|||||
Db 2700 TTGTAGGAGTCCCTCAGCACAGCTGCCCTCTCAGATATAGAAAGATCCTTGAGACTC 2759
Qy 156 LysThrLeuGluArgGluIleSerGlyGlyAlaMetAlaValAlaValIleLeuLeuAan 175
2760 AAGACGTTAGAGAGGAAATTTGCGGAGGGGCGCATGCGCTGTGTGCGGCTTCTCAAC 2819
Qy 176 AsnLysLeuIyrValAlaAanVal----- 183
Db 2820 AACAACTTACGTCCCAATGTGTGAGCCCTCTGTCTCCAGGACAGGAGGACTG 2879
Qy 183 ----- 183
Db 2880 GGGAGAGTCAAGCCACAGGGGTGGTGATTAATTGACAATCTGCTTCCAGACACTTCA 2939

QY 183 ----- 183
Db 2940 CGCACTTTAAACCCAGGGTCTCCTGAGACGGTTGGGTATGTCCCTCTCCACAGTGAGCC 2999
QY 183 ----- 183
Db 3000 TCAGTCCCAACTGGAAGGAGAAAGACGGGATGGGAGACAGGTGCTCGCCCTTTAGTC 3059
QY 183 ----- 183
Db 3060 CCTATTCTGCTTCTTAACCTCAITCTTGACAGTCGTAATCCCATTTGGGTGGCTTTGGGG 3119
QY 183 ----- 183
Db 3120 GCCGCTGCTTCTGGAAGAGGTTACCTGGCCATGAATAACCAACCACTATCCCATGAACA 3179
QY 183 ----- 183
Db 3180 GCTCGAGATGGGCTAGGTGACAGGGACATTTGGGGTTTTGTGAGAAGACCAGACAGGTGC 3239
QY 183 ----- 183
Db 3240 AGGTTTCAGTAGAAGGACTCTGTAGAGACCCTTCTGATGATGCTGCCTTTTTTAAATAC 3299
QY 183 ----- 183
Db 3300 TCCAGTGGAAATTTTCCACAATATAGAACAAATAGAGTGACTGATATTTAATGAACATTC 3359
QY 183 ----- 183
Db 3360 ATGGGCCCATCCCCAATCCAGCAGTATCAACTGTGGCGAGCCTTCTTTGTCCCCACGT 3419
QY 183 ----- 183
Db 3420 TCTCTATTTGAAGCAATCCAGACACTGTATCATGTTATCTGTAATATTTCACTCTG 3479
QY 183 ----- 183
Db 3480 TATCTCTAGAGAGAGTCTGTTTTATTAGATCATAACTCTATGATTACACTGAAAGAG 3539
QY 183 ----- 183
Db 3540 TTAAACCCAAATCTTTCGTATCACCACTGTCCCTAACTGGTGACTCACCCCGTTGGTCT 3599
QY 183 ----- 183
Db 3600 GAGCCTGTTTGGCCATTTTCAGGTATTTCCATGTGTGAATGCTGCTTTTCCCTCTCT 3659
QY 183 ----- 183
Db 3660 GCCTTCCGGTATGCCCTATTTCTCTGTGTAGTCTTTGCTTAGCTGTTCCACATTCT 3719
QY 184 -----GlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThr 201
Db 3720 GCCACAGGTACAAACCGTGACATTTATGCAATCGACAGTGGATGGTTGCAGGTGACA 3779
QY 202 GlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeu 221
Db 3780 CAGCTGAACGTGGACCAACACCAGAGAACGAGGATGAGCTCTTCGCTCTTTTCGCAGCTG 3839
QY 221 ----- 221
Db 3840 GGTGAGTGGGAGATGGGAGCGGAAGCTGATCCCCATGGGCTCACCTTCGCTCGCTGCTT 3899
QY 221 ----- 221
Db 3900 TGGTGGTGGGTAGAGAGCGGTGTGTGAGAGGGCTGTGATCTTGGGCTCCCCAGCCAGC 3959
QY 221 ----- 221
Db 3960 CTGCTGGGGTTTCAATCCCAGCACTGCCGCTTACTGGTTGGTGATTGGAACTTTAAGTTTT 4019

QY 221 ----- 221
Db 4020 CTGTGCTTCAGATCCCGCTGTGTAAATGAGGATAATAGTAACCTACCTCGGGATGCTAG 4079
QY 221 ----- 221
Db 4080 GAGGATTCAGTTAGTGCATGTGAAATGCTTCTTGAGTGCCTGGCACACAGAAGACACTT 4139
QY 221 ----- 221
Db 4140 ACTGTTATTGTGGAGACTGAAAGAGGCCAAAGATCCAGGAGGCCAGCTGCTGTCTGA 4199
QY 221 ----- 221
Db 4200 GCTGCTCCTTCCCTTCCAGTGAGCTCCAGGCAGATGTGGGGCACACGGGATGGCAGGG 4259
QY 221 ----- 221
Db 4260 AGACGGCAGGCAAGCTGCTCCGTGCCAGTGTGCTCTGGAGACGGGCTCTGAGTGAGGT 4319
QY 221 ----- 221
Db 4320 GTGTCGTATGGCGTGGGACTGAGGACACCAGGAGCTTTGGGTGAGTCTCTCCAG 4379
QY 221 ----- 221
Db 4380 GTGCCCTGGTGTGCTTCTTCAATTCATTCCAGTGGGTCTTTATTGCTTCTTCCCATGAC 4439
QY 221 ----- 221
Db 4440 TGTGCTCTGTCCCTTCTTTTCTTCTTCTTGTGAACAAGACAGGATTGTTGCAC 4499
QY 222 -----GlyLeuAspAlaGlyLysValLysGlnValGlyLileCysGlyG 237
Db 4500 GTTTCCTCCGTAGCTTGGATGCTGGAAGATCAAGCAGGTGGGATCATCTGTGGCA 4559
QY 237 nGluSerThrArgArgLileGlyAspTyrLysValLysTyrGlyTyrThrAspLileAspLe 257
Db 4560 GGAGAGCACCCCGCGGATTCGGGATTACAAGTTTAAATATGGCTACACGACATTGACCT 4619
QY 257 uLeuSer-----AlaAlaLysSerLysProLileAlaGluProGluLileHisGlyAl 275
Db 4620 TCTCAGTAGTAGTGCCAGCCAGCTGTCCCTGTGCTTGAAGAACAAGAA----- 4668
QY 275 aGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGly-LeuTyrLysA 295
Db 4669 -GGTCCTAGGAGGCCAAGATGGGAGGATTGATTGTGCTG-----GGAGTCGAGGCTG 4721
QY 295 laLeuGluAlaAlaHisGly-----ProGlyGlnAlaAsnGlnGluileAla- 310
Db 4722 CAGTGAGCCGTGATCGCCACTGCACCTCCAGCCTGGGCAACAGCAAGACCCCTGTCTC 4781
QY 311 -----AlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaG 328
Db 4782 AAAAAAAAAAAGTCCAGAAAGGCTGGGCGACAGC---AGACTGTAGGCCCCAGTGA 4838
QY 328 lnAlaValValAspArgValLysArgLileHisSerAspThrPheAlaSerGlyGlyGluA 348
Db 4839 CACATGTGCACTCAAGTAGTTCAGAGCCCGAGCAAACTCTTTCTTCTCTC---CCAGAGA 4895
QY 348 rg-----AlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsn 365
Db 4896 GGTGTTTGGGGAGGTTTAGGCCCAGCAGCCAGGAGTTGGGTACTGTAGAAACAAGGG 4955
QY 365 heGlyTyrProLeuGlyGluMet-----SerGlnProThrPro-----S 378
Db 4956 GTGGAGTTGAGTGGGCCAAAGCTGAGGGCAGCAACTAGTCCCAAGGGAGGAGCGGCA 5015
QY 378 erProAlaProAlaAlaGlyGlyArgValTyrProValSer-----ValProTyrSerS 396
Db 5016 AGCCAGAGTTGGATGGGATGGACAGCACAGCCAGCCTGTGGTGGTGTGCCCATGGCTGC 5075
QY 396 erAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetProSerGln---- 414

Db 5076 AGCAGAG-----TGTGATGATGATTGTATTTCCCTCTTGAGGACA 5117
Qy 415 -----GlyGlnMetValasnGlyAlaHisSerA 424
Db 5118 CCTCTCTGTGACATCACCCAGCGCTCATGGCAGCAAGATGGGACGATCATCACACGG 5177
Qy 424 laseThrIleuAspGluAlaIatnProThrIleuThraSngInserProthrIleuThraLeug 444
Db 5178 CCAAGCCCGGCGAGAGCTGTGATGTACTCTTTTATGCAATTCCTACTCAGGCCCCCT 5237
Qy 444 lnsertHraSnthRhs-ThrgInserSerSerSerSerAspGlyGlyIleuPhary 463
Db 5238 GAGGCGATCTGCTCCATCACCCAGAGGCATTTTTCATTAAGAGGAACGTAGGCATGGAGA 5297
Qy 464 SerArgProAlaHisSerIleuProGly 473
Db 5298 CATCAGGCAGCCAC-----CCTGGT 5318
RESULT 9
AAL36984
ID AAL36984 standard; DNA; 16877 BP.
XX
AC AAL36984;
XX
DT 08-JAN-2002 (first entry)
XX
DE Human musculoskeletal system related polynucleotide seq ID NO 3349.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antitiral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticulcer;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein;
KW musculoskeletal system; ds.
XX
OS Homo sapiens.
XX
PN WO200155367-A1.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01338.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205115.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.

PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246538.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.

PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 11-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
PA (HUMA-) HUMAN GENOME SCI INC.
XX

Rosen CA, Barash SC, Ruben SM;
WPI; 2001-451937/48.

Isolated polypeptide for treating, preventing and/ or prognosing
disorders related to the musculoskeletal system including
musculoskeletal cancers and also for testing and detection e.g.
diagnosis -

Example 2; SEQ ID NO 3349; 781pp + Sequence Listing; English.

The invention relates to novel genes (AAL34669-AAL37666) and proteins (ABB03087-ABB04109) associated with the musculoskeletal system useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemia; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from wipo at ftp.wipo.int/pub/published_pct_sequences.

Sequence 16877 BP; 3563 A; 4508 C; 4516 G; 4290 T; 0 other;

Alignment Scores:

Pred. No.: 2,27e-30 Length: 16877
Score: 543.50 Matches: 295

Percent Similarity: 18.86% Conservative: 43
Best Local Similarity: 16.46% Mismatches: 107
Query Match: 21.07% Indels: 1348
DB: 22 Gaps: 19

US-09-830-144-4 (1-504) x AAL36984 (1-16877)

QY 12 GluGlnGlnProSerTyrThrAspAspLeuProLeuCysHisLeuSerGlyValGlySer 31
Db 1 GAGCAGCAGCCAAAGCTGGACAGATGCTCTCTGCCACCTCTCTGGGGTTGGCTCA 60
QY 32 AlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisProGluAsp 51
Db 61 GCCTCCAAACCGCAGCTACTCTGCTGATGGCAAGGSCACTGAGAGCCACCCCGCAGGAC 120
QY 52 SerTyrLeuLysPhe----- 56
Db 121 AGCTGGCTCAAGTTTCAGGTGTGTGTGCCAGCATTTCTGTGTGGCCCGGGAGTTGGTT 180
QY 56 ----- 56
Db 181 GGTTTGCAAGCAAGAAAGACACCGACCTTCGAGCTTTCTCGTATGGGCTTGCAGTGAC 240
QY 56 ----- 56
Db 241 ATGTGGCCCGTGAGAGGTGGCTCTGCTGCTGTCTTGCCAAAGGGCCTGCTCTGATGGGT 300
QY 56 ----- 56
Db 301 AGCGTGAGCATGGGAGGAGGGGTATCCCAAGATGTCTATAGCCAGAGTGAATGATGGCTA 360
QY 56 ----- 56
Db 361 AAGCAGGGGGACCCAGGAGGGCCCTGAAGCTGCAGCTGTGTCGTTTAGTCTCCCCCA 420
QY 56 ----- 56
Db 421 ATTCCTTTTCCCTTCCCTCCACCTCCGTGAGACCCCTGGTCTCAGGCCCTCCCTCTGCC 480
QY 57 -----ArgSerGluAsnAsnCysPheLeuTyrGlyValPheAsnGlyTyrAsp 72
Db 481 TCTCCCTCTCCAGGAGTGAGAACAACTGCTTCTGTATGGGGTCTTCAACGGCTATGAT 540
QY 73 GlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAlaGluLeuLeuGlyGln 92
Db 541 GGCACCGAGTGACCACTTCTGGCCAGCGGCTGTCCGACAGAGCTCTGCTGGGCCAG 600
QY 93 LeuAsnAlaGluHisAlaGluAlaAspValArgArgValLeuLeuGln----- 108
Db 601 CTGAATGCCGAGCAGCGCGAGCGGCGATGTGGGGCGTGTGCTGTGTCAGGTAATGGTCCG 660
QY 108 ----- 108
Db 661 GGGCCAAAGTGAACCCAGCCACATATGTCCCCACCCCAAGGCTTGGGCCCTGCAGCTC 720
QY 108 ----- 108
Db 721 TAGCATGTTGCCAGGGTTGTTGTAAGATCTCGCGGCCCCCTTCAACCCAGTAGAGGAGC 780
QY 108 ----- 108
Db 781 AGCTCCAGCGTAGGCCCCCCCCCAACAGGAGTCCAGGACCCAGCCAGCAGCAGG 840
QY 108 ----- 108
Db 841 CTTTCAATGGCTGGCAGAGGAGGTACTCTGAGACGGGGGATTTAGGATGGAGCTT 900
QY 108 ----- 108
Db 901 GGAGAGAGGTGTGAGGTGGGAGCAGGGCAAGCCCTGTAGAAATGGGGTCATTTAGAGCT 960
QY 108 ----- 108

Db 961 ACCCCTTTCTTCTATGTGTGTCAGGTGCTCAGGTCAGAGGACCATCCCTGG 1020
QY 108 ----- 108
Db 1021 CGTCAGAGAGACCTTGCCCTCCCTTTCTGAGGGCCGCCCTCATTTGACTGTTCC 1080
QY 108 ----- 108
Db 1081 ACCTTCTGTGCTGCTGTGGCAGTTTTCCTGTGGCATCTGCTTTGAGAGCATGTCT 1140
QY 108 ----- 108
Db 1141 CAGGCCCATTTGATGAGAAGATGGCTTCTGTTCCGGAGAGGGGTGTCCAGCCTT 1200
QY 108 ----- 108
Db 1201 TTCTGTCCCTTCAGACACTCAGGCTTCATTGCGAGTATCTCAGCAGATCTCAGCAGG 1260
QY 108 ----- 108
Db 1261 GGGAGAGGTGTCAGTGTCAACAGTGTCTGTGGGCTGTGGGTTTGACAGAACTCTCC 1320
QY 108 ----- 108
Db 1321 AGAAGAAGTTGTAGTAATCATACAGCTGTCAATAGGCTGTGGCAGTTTCTGTGACAGTT 1380
QY 108 ----- 108
Db 1381 GCCTTGCTGAGTGCAGGTTTGA CTGTCTTACCTAACCCCTGGAACAACACTTAACCTCC 1440
QY 108 ----- 108
Db 1441 GTTGTGTGAGACTGAGGGGCGCAGAGTCAACACAGCTGGGCTTACGCCAACCTTTGCT 1500
QY 108 ----- 108
Db 1501 CTTAGCTTCGCACTGTCTCTGCTCCAGCCGCTGTGCAAGTGTGCTCTGTGAGGGTGC 1560
QY 108 ----- 108
Db 1561 ACACCCCTTCGAGGCTGAAGGGCTTTGTCAAGACATTGATCTGCAGAGACCGGTG 1620
QY 108 ----- 108
Db 1621 CCTTGAGTGTGCGCCAGAGCAGACTGACATGTGAAGCTCCATCACAGAAACC 1680
QY 108 ----- 108
Db 1681 TGCAGTGAAAGACAGAAAGCTGCTGTGATTAAATAGAGACATTTTGGCACAGTGAC 1740
QY 109 ----- 113
Db 1741 AGTGTGTTTTGAACCAAGCTTTGCCCTGTCTGTGTCCCTCAGGCTTCGATGTGGTG 1800
QY 114 GluArgSerPheLeuGlnSerIleAspAspAlaLeuAlaGlnIleLysAlaSerLeuGlnSer 133
Db 1801 GAGAGGAGCTTCTCGGATTCATTGACGACGCTTGTGCTGAGAAAGCAAGCCTCCAGTGG 1860
QY 134 GlnLeuProGlu----- 137
Db 1861 CAATTGGCAGA-GGTAAATTTCCCGAGCCGACCCAGGGAGTCAAGTCCAGGCCCAAGCT 1919
QY 137 ----- 137
Db 1920 TTGCAAGAGACATGACTCATCTACTTTCTTGACATTACTGGGCCAGAGCAACAGCGTT 1979
QY 137 ----- 137
Db 1980 AGGAGCAGTTCTCTGATGGGTGACACTGTGTGTGGCACAGGTGAGGACCTGTGCT 2039
QY 137 ----- 137
Db 2040 CTGCTGTGGCAGCTGCATGCTAAAGCTCAGCAGTGTCTGGGACGTGTGGACATGAGGA 2099

QY 137 ----- 137
Db 2100 AGGCTTTTCCAGTCTCTGTCTGTGAATGGTCCCGGCCCTTTGTCAACCACTGCTCAC 2159
QY 137 ----- 137
Db 2160 ATGGCTACTGCACCGTCGACGGCTTACCTGTGTGCTGTGAGTGTGGTATCTCCAGCTT 2219
QY 137 ----- 137
Db 2220 TCTGAAGAACTTAGCTTGAGCTGTAGCGAAGCTTCATCTTGCTGTGCCATGAGACC 2279
QY 137 ----- 137
Db 2280 ATCTGTCTGTGCTCCCGGGCCCTCGGATCTTTGCTCTTAAGACAAAGGAGGCCGG 2339
QY 137 ----- 137
Db 2340 GCATGTGTGCTCAGCGCTGTAAATCCAGACACTTGGAGAGCCGAGGCGGATCACA 2399
QY 137 ----- 137
Db 2400 GGTGAGAGATCAAGACCATCTGGCTAACATGTGAAAACCCGCTCTTAATAAATAC 2459
QY 137 ----- 137
Db 2460 AAAAAATTAGCTGGGCGTGTGTGGGCGCTTTGTAGTCCAGCTACTCTCGAGGCTGAA 2519
QY 137 ----- 137
Db 2520 GCAGAGAAATGTGTGAACCTGGGAGCGGAGCTTGTAGTGAACGAGATCATGCCACTG 2579
QY 137 ----- 137
Db 2580 CACTGTAGCTGGGGGACACAGGAGACTCATCTGAAAAAAGTAAAGG 2639
QY 137 ----- 137
Db 2640 AGTGAAGGGCTAAGAAATGCTCCAGAGTTGTGTCTGCTGACCTCTGTATGG 2699
QY 138 ----- 155
Db 2700 TTGTAGGAGTCCCTCAGCACACACTGCTCTCAGTATAGAAAGATCCTTGAGAGCTC 2759
QY 156 LysThrLeuGlnArgGlnIleSerGlyAlaMetAlaValAlaValLeuLeuAsn 175
Db 2760 AAGACGTTAGAGAGGAAATTTGCGGAGGGCCATGGCCTGTGTGGGGTCTTCTCAAC 2819
QY 176 AsnLysLeuIyrValAlaAsnVal----- 183
Db 2820 AACAACTTAACGTTCGCCAATGTGTGAGCCCTCTGTGCCAGGGCAGGAGAGCTG 2879
QY 183 ----- 183
Db 2880 GGGAGAGTCAAGCCACAGGGGTGCGTGCATTATTGCAATCTGCTTTCCAGACACTTCA 2939
QY 183 ----- 183
Db 2940 CGCACTTAAACCAAGGTCTCTGAGACCGTTGGTATGTCCCTTCACAGTGAAGCC 2999
QY 183 ----- 183
Db 3000 TCAGTCCCACTGGAAGGAGAAAGACGAGATGGAGACAGTGTCTGGCTTTAGTC 3059
QY 183 ----- 183
Db 3060 CCTATTCTGCTTTTAACATCTTGTGACAGAGTGTATCCCATTTGTGGGTGGCTTGGGG 3119
QY 183 ----- 183
Db 3120 GCCGCTGCTTGTGAAGAGTTTACTGTGGCATGAATAACCAACAGTATCCCATGAACA 3179

QY 183 ----- 183
Db 3180 GGTGAGATGGGGCTAGGTGACAGGACATTTGGGGTTTGTGAGAGACAGAGTGC 3239
QY 183 ----- 183
Db 3240 AGGTTTCAGTAGAAGGACTCTGTAGAGACCTTCTGATGATGCTGCTTTTAAATAC 3299
QY 183 ----- 183
Db 3300 TCCAGTGGAAAAATTTCCACAATATAGAACAAATAGAGTGAATGATATTAATGAACATTC 3359
QY 183 ----- 183
Db 3360 ATGCCCCATCCCAATCCAGCAGTTATCAACTGTGGCCAGCCCTCTTTGTCCCCACGT 3419
QY 183 ----- 183
Db 3420 TCTCTATTTGAAGCAATGCCAGACACTGTATCATGTTATCTGTAATAATTTTCAGTCTG 3479
QY 183 ----- 183
Db 3480 TATCTCTAGAGAAGAGTCTGTTTATTAAAGATCATATCCTATGATTAACACTGAAAAAG 3539
QY 183 ----- 183
Db 3540 TTAAACCCCAATTCCTGATACACAGCTGTCCCTAACTGGTGACTCACCCCGTTGGTCT 3599
QY 183 ----- 183
Db 3600 GAGCCTGTTTGGCCCAATTCAGGTATTTCCATGTGTGAAATGCTGCTTTTCCCTCTCT 3659
QY 183 ----- 183
Db 3660 GCCTCCCGGTATGCCCTATTCTCTGTGTGTAGTCTTTGTGTAGCTGTTCACATTCT 3719
QY 184 ----- GlyThrAsnArgAlaLeuLeuCysLysSerThrValaspGlyLeuGlnValThr 201
Db 3720 GCCACAGGTACAAACCGTGCACTTTTATGCAATCGACAGTGGATGGTTGCAGGTGACA 3779
QY 202 GlnLeuAsnValAspHisThrThrClnuGlnuAspGluLeuPheArgLeuSerGlnLeu 221
Db 3780 CAGCTGAACGTGGACACACACAGAGAACAGGATGAGCTCTTCCTGCTTTCGCAGCTG 3839
QY 221 ----- 221
Db 3840 GGTGAGTGGGAGAGTGGGAGCGGAAGCTGATCCCCATGGGCTCACCTTCGCTGCTT 3899
QY 221 ----- 221
Db 3900 TGGTGTGGGTAGAGAGCGGTGTGTAGAGGGGCTGTGATCTTGGGCTCCCCAGCCAGC 3959
QY 221 ----- 221
Db 3960 CTGCTGGGGTTCAATCCAGCAGCTGCCGCTTACTGGTTGTGATTGGAATTAAGTTTT 4019
QY 221 ----- 221
Db 4020 CTGTGCTCAGATCCCCCGCTGTGTAATGAGGATAATAGTAACCTACCTCGGGATGCTAG 4079
QY 221 ----- 221
Db 4080 GAGGATTCAGTTAGTGCAATGTGAAATGCTTCTGGAGTGCCTGGGCACAGAGACACTT 4139
QY 221 ----- 221
Db 4140 ACTGTTATTGGTGAGACTGAAAGAGGCCAAAGAGTCCAGGAGCCAGCTGCTGCTGA 4199
QY 221 ----- 221
Db 4200 GCTGCTCCTTCCCTTCCAGTGAGCTCCAGGCAGATGTGGGGCACACGGGATGGCAGGG 4259
QY 221 ----- 221

Db 4260 AGACGGCAGCAGCAAGCTGCTCCGTGCCAGGTGGTGCTCTGGAGGACGGGCTCTGAGTGAGGT 4319
QY 221 ----- 221
Db 4320 GTGTCGTGATGGGCGTGGGACTGAGGACACACAGGAGCTTTGGGTCAGCTGCTCTCCAG 4379
QY 221 ----- 221
Db 4380 GTGCCCTGGTGTGTCTTCAATTTCTATTCAAGTGGGTCTTATTGCTCTTCTCCCATGAC 4439
QY 221 ----- 221
Db 4440 TGTGTCTCTGTCCCTTCTTTTGTTCCTCTTTGTGAACAAGAGCAGGATGTTTGCACCT 4499
QY 222 ----- GlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCysGlyGln 237
Db 4500 GTTTCCTCCGTAGGCTTGGATGCTGGAAGATCAAGCAGTGGGATCATCTGTGGGCA 4559
QY 237 nGluSerThrArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLe 257
Db 4560 GGAGAGCACCCCGCGGATCGGGGATTACAAGGTTAAATATGGCTACACGGACATTGACCT 4619
QY 257 uLeuSer ----- AlaAlaLysSerLysProIleIleAlaGluProGluIleHisGlyAl 275
Db 4620 TCTCAGGTAGGTGCGCAGCCAGCTGCTCCCTGTGCTTGAAAGAACAGAA ----- 4668
QY 275 aGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSerGluGly-LeuTyrLysA 295
Db 4669 -GGTCTAGGAGCCCAAGATGGGAGGATTCATTTGCTG ----- GGAGTCGAGGCTG 4721
QY 295 laLeuGluAlaAlaHisGly ----- ProGlyGlnAlaAsnGlnGluIleAla- 310
Db 4722 CAGTGAGCGGTGATCGCCACTGCACTCCAGCTGGGCAACAGAGCAGACCTGTCTC 4781
QY 311 ----- AlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaG 328
Db 4782 AAAAAAAGGTCCAGAGCCCTGGGCGCAGACG ----- AGACTGTAGGCCCACTGCA 4838
QY 328 lnAlaValValAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGlu 348
Db 4839 CACATGTCAGCTCAAGTAGTTTCAGAGCCAGGCAAACTCTTTCTCTCTC ----- CAGAGA 4895
QY 348 rg ----- AlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsn 365
Db 4896 GGTGTTTGGGCGAGGTAGGCCAGCAGCAGGAGGTGGGTACTGTAGAAAAAAGGG 4955
QY 365 heGlyTyrProLeuGlyClnuMet ----- SerGlnProThrPro ----- S 378
Db 4956 GTGGAGTTGAGCTGGGCCAAAGCTGAGGGCAGCACTAGTCCCAAGGAGGAGCGCGCA 5015
QY 378 erProAlaProAlaAlaGlyArgValTyrProValSer ----- ValProTyrSerS 396
Db 5016 AGCCAGAGTTGGATGGATGGACAGACAGGCTGTGTGTCGGTGTGCTGCCATGGCTGC 5075
QY 396 erAlaGlnSerThrSerLysThrSerValThrLeuSerLeuValMetProSerGln ----- 414
Db 5076 AGCAGAGG ----- TGTGAGATGAATGTTGTTATTTCCCTCTGAGGACA 5117
QY 415 ----- GlyGlnMetValAsnGlyAlaHisSerA 424
Db 5118 CCTCTCTGGACATCACCCAGCGCTCATGGCAGCAATGGGCGAGTCATCATCACGG 5177
QY 424 laSerThrLeuAspGluAlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuG 444
Db 5178 CCAAGCCGTCGAGAGCTCTGGATGTACCTCTCTTTTATGCAATTCCTACTCAGGCCCTC 5237
QY 444 lnSerThrAsnThrHis-ThrGlnSerSerSerSerSerSerSerSerSerSerSerSer 463
Db 5238 GAGGAGTCACTGCTGCTCAGCCAGAGGCAATTTTCAAGAGGAACTAGGAGCATGGAGA 5297
QY 464 SerArgProAlaHisSerLeuProProGly 473
QY 473 -----

Db 5298 CATCAGCAGCCAC-----CTGGT 5318

RESULT 10

AAFI5895

ID AAFI5895 standard; cDNA; 696 BP.

XX

AC AAFI5895;

XX

DT 13-MAR-2001 (first entry)

XX

DE Human prostate cancer antigen nucleotide sequence SEQ ID NO:330.

XX

KM Human; prostate cancer; prostate cancer antigen; detection; diagnosis;

KM neuroprotective; cytosolic; cardioactive; immunomodulatory; muscular;

KM vulnerrary; gastrointestinal; nephrotoxic; antiinfective; gynaecological;

KM antibacterial; gene therapy; neural; immune; reproductive; renal;

KM gastrointestinal; pulmonary; cardiovascular; proliferative disorder;

KM wound; infectious disease; ss.

XX

OS Homo sapiens.

XX

PN WO200055174-A1.

XX

PD 21-SEP-2000.

XX

PF 08-MAR-2000; 2000WO-US05988.

XX

PR 12-MAR-1999; 99US-0124270.

XX

PA (HUMA-) HUMAN GENOME SCT INC.

PA (ROSE/) ROSEN C A.

XX

PI Rosen CA, Ruben SM;

XX

DR WPI, 2000-587513/55.

XX

DR P-PSDB; AAB56692.

XX

PT Prostate cancer associated gene sequences, referred to as prostate

PT cancer antigens, useful for treatment, prevention, and diagnosis of

PT disorders such as prostate cancer -

XX

XX

PS Claim 1; Page 837; 2338pp; English.

XX

CC AAFI566 to AAFI605 encode the human prostate cancer associated

CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.

CC The prostate cancer antigens can have neuroprotective, cytosolic,

CC cardioactive, immunomodulatory, muscular, vulnerrary, gastrointestinal,

CC nephrotoxic, antiinfective, gynaecological and antibacterial activities,

CC and can be used in gene therapy. The prostate cancer antigen

CC polynucleotides may be used for detection of prostate cancer, chromosome

CC identification, as chromosome markers, and for numerous other diagnostic

CC or research purposes. The prostate cancer antigens may be used to treat

CC disorders such as neural, immune, muscular, reproductive,

CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative

CC disorders, wounds, and infectious diseases. AAFI606 to AAFI6514 to

CC AAB57303 represent sequences used in the exemplification of the present

CC invention.

XX

SQ Sequence 696 BP; 143 A; 237 C; 187 G; 120 T; 9 other;

XX

Alignment Scores:

Pred. No.: 5.27e-22 Length: 696

Score: 409.00 Matches: 78

Percent Similarity: 98.75% Conservative: 1

Best Local Similarity: 97.50% Mismatches: 1

Query Match: 15.85% Indels: 0

DB: 21 Gaps: 0

US-09-830-144-4 (1-504) x AAFI5895 (1-696)

OY 425 SerThreuspglualThrProThleuThrAsnGlnSerProThrleuThrLeuGln 444

Db 13 GCCACCTGGAGAGCAGCCACCCCTCAACCAACCAAGCCGACCTTAACTCGAG 72

OY 445 SerThraanthThiThrglnSerSerSerSerSerAaspGlyGlyLeuPheArgSer 464

Db 73 TCACCCACACACACACAGCAGCAGCTCCAGCTTACGAGAGGCTCTCCGCTCC 132

OY 465 ArgProAlaHisSerLeuProProGlyGlnAaspGlyArgValGluProTyrValAaspPhe 484

Db 133 CGGCCCGCCCATCGCTCCGCTCGCAGAGACGGTCCGTGTGAGCCCTTAGTGACTTT 192

OY 485 AlaGluPheTyrArgLeuTyrPserValAaspHisGlyGlnSerValValThrAlaPro 504

Db 193 GCTGAGTTTACCGCCTCGAGCGTGCATGCGCAGACAGCGTGTGACACACCG 252

RESULT 11

ABAA7246

ID ABAA7246 standard; DNA; 211 BP.

XX

AC ABAA7246;

XX

DT 01-FEB-2002 (first entry)

XX

DE Human breast cell single exon nucleic acid probe #5941.

XX

KM Human; microarray; single exon probe; gene expression; breast;

KM disease; cancer; ss.

XX

OS Homo sapiens.

XX

PN WO200157271-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JUN-2001; 2001WO-US00662.

XX

PR 04-FEB-2000; 2000US-0180312.

XX

PR 26-MAY-2000; 2000US-0207456.

XX

PR 30-JUN-2000; 2000US-0608408.

XX

PR 03-AUG-2000; 2000US-0632366.

XX

PR 21-SEP-2000; 2000US-0234687.

XX

PR 27-SEP-2000; 2000US-0236359.

XX

PR 04-OCT-2000; 2000GB-0024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-496933/54.

XX

PT New spatially-addressable set of single exon nucleic acid probes,

PT useful for measuring gene expression in sample derived from human

PT breast, comprises number of single exon nucleic acid probes -

XX

PS Claim 4; SEQ ID NO 5941; 327bp + sequence listing; English.

XX

CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human breast and BT 474 cells. The method involves contacting

CC the probes with a collection of detectably labelled nucleic acids

CC derived from mRNA of human breast, and then measuring the label

CC bound to each probe of the microarray. The probes are useful for

CC verifying the expression of regions of genomic DNA predicted to

CC encode proteins. They are useful for gene discovery, and for

CC determining predisposition and/or prognosing breast disease. Gene

CC expression analysis is useful for assessing the toxicity of chemical

CC agents on cells. The microarray of this invention presents a far greater

CC diversity of probes for measuring gene expression, with far less bias

CC than expressed sequence tag microarrays. The method is suitable for

CC rapid production of functional information from genomic sequence. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;
Alignment Scores:
Pred. No.: 2.38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0

US-09-830-144-4 (1-504) x ABA47246 (1-211)

QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCACAGCAGACCTCCCTGGACGAGTGGCCAGCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGTGAAGCCATCCACAGCAGACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGCAGCAGGACATGACCTCTAGTGAGGAACCTTTGGCTACCGCTGGCGGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCAGCCCCA 210

RESULT 12

ABA65131

ID ABA65131 standard; DNA; 211 BP.

XX AC ABA65131;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #13436.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-483447/52.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human fetal liver -

XX PS Claim 4; SEQ ID NO 13436; 639pp + sequence listing; English.

XX XX The invention relates to a single exon nucleic acid probe for

XX CC measuring human gene expression in a sample derived from human foetal

XX CC liver. The single exon nucleic acid probes may be used for predicting,

XX CC measuring and displaying gene expression in samples derived from human

XX CC foetal liver. The present sequence is a single exon nucleic acid

XX CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;
Alignment Scores:

Pred. No.: 2.38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0

US-09-830-144-4 (1-504) x ABA65131 (1-211)

QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCACAGCAGACCTCCCTGGACGAGTGGCCAGCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGTGAAGCCATCCACAGCAGACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGCAGCAGGACATGACCTCTAGTGAGGAACCTTTGGCTACCGCTGGCGGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCAGCCCCA 210

RESULT 13

ABA32233

ID ABA32233 standard; DNA; 211 BP.

XX AC ABA32233;

XX DT 23-JAN-2002 (first entry)

XX DE Probe #10699 for gene expression analysis in human heart cell sample.

XX KW Human; gene expression; heart; microarray; vascular system; probe;

XX KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX KW congenital heart disease; ss.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX WPI; 2001-488899/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT hearts -

XX PS Claim 4; SEQ ID No 10699; 530pp; English.

XX XX

CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart. The
CC present sequence is one such probe. The probes may be used for
CC predicting, measuring and displaying gene expression in samples derived
CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;

Alignment Scores:

Pred. No.:	2,38e-19	Length:	211
Score:	365.00	Matches:	70
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	14.15%	Indels:	0
DB:	22	Gaps:	0

US-09-830-144-4 (1-504) x ABA32233 (1-211)

OY 312 MetIleAspThrGluPheAlaIysGlnThrSerLeuAspAlaValAlaGlnAlaVal 331

DB 1 ATGATTGACACTGAGTTTGCCAAAGCAGACCTCCCTGGAGCGAGTGCCAGGCCGTCTG 60

OY 332 AsparGValIysArgIleHisSerAspThrPheAlaSerGlyGlyUargAlaArgPhe 351

DB 61 GACCGGGTGAAGCGCATCCACAGCCACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120

OY 352 CysProArgHisGluAspMetThrLeuLeuValArgAsnPhgIlyTyProLeuGlyGlu 371

DB 121 TGCCCCCGGACAGAGCATGACCTGCTAGTGAAGAACTTGCTACCCGCTGGCGAA 180

OY 372 MetSerGlnProThrProSerProAlaPro 381

DB 181 ATGAGCCAGCCACACCGAGCCAGCCCA 210

RESULT 14

AAK13550
ID AAK13550 standard; DNA; 211 BP.

XX AAK13550;

DT 05-NOV-2001 (first entry)

XX Human brain expressed single exon probe SEQ ID NO: 13541.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer; ss.

XX Homo sapiens.

PN WO200157275-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00667.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX Single exon nucleic acid probes for analyzing gene expression in human

PT brains -

PS Example 4; SEQ ID NO: 13541; 650bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is one of the probes of the

CC invention.

XX Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;

SQ US-09-830-144-4 (1-504) x AAK13550 (1-211)

OY 312 MetIleAspThrGluPheAlaIysGlnThrSerLeuAspAlaValAlaGlnAlaVal 331

DB 1 ATGATTGACACTGAGTTTGCCAAAGCAGACCTCCCTGGAGCGAGTGCCAGGCCGTCTG 60

OY 332 AsparGValIysArgIleHisSerAspThrPheAlaSerGlyGlyUargAlaArgPhe 351

DB 61 GACCGGGTGAAGCGCATCCACAGCCACCTTCGCCAGTGGTGGGAGCGTGCAGGTTT 120

OY 352 CysProArgHisGluAspMetThrLeuLeuValArgAsnPhgIlyTyProLeuGlyGlu 371

DB 121 TGCCCCCGGACAGAGCATGACCTGCTAGTGAAGAACTTGCTACCCGCTGGCGAA 180

OY 372 MetSerGlnProThrProSerProAlaPro 381

DB 181 ATGAGCCAGCCACACCGAGCCAGCCCA 210

RESULT 15

AAK39289
ID AAK39289 standard; DNA; 211 BP.

XX AAK39289;

DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 13846.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

PN WO200157276-A2.

PD 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.

PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.


```
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488900/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
XX Example 4; SEQ ID NO: 13846; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
XX Sequence 211 BP; 43 A; 70 C; 66 G; 32 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2,38e-19 Length: 211
Score: 365.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.15% Indels: 0
DB: 22 Gaps: 0
US-09-830-144-4 (1-504) x AAK39289 (1-211)
QY 312 MetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaValVal 331
Db 1 ATGATTGACACTGAGTTTGCCAGACAGACCTCCTGGACGACGTGCCCGCCGTCGTG 60
QY 332 AspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAlaArgPhe 351
Db 61 GACCGGGTGAAGCGCATCCACAGCGACACCTTCGCCAGTGTGGGAGCGTCCAGGTTTC 120
QY 352 CysProArgHisGluAspMetThrLeuValArgAsnPheGlyTyrProLeuGlyGlu 371
Db 121 TGCCCCCGGCACGAGGACATGACCTGCTAGTGAGGAACCTTTGGCTACCCGCTGGCGAA 180
QY 372 MetSerGlnProThrProSerProAlaPro 381
Db 181 ATGAGCCAGCCACACCGAGCCCGCCCA 210
```

Search completed: December 9, 2002, 23:14:45
Job time : 284.44 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2002 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 9, 2002, 22:59:44 ; Search time 1573.89 Seconds
(without alignments)
5186.195 Million cell updates/sec

Title: US-09-830-144-4

Perfect score: 2580

Sequence: 1 MAAQRRLSLQSQPQSWTDD.....AEFYRLWSVDHGQSVVTAP 504

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool/US09830144/runat_04122002_141353_2264/app_query.fasta_1.1422
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -DOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09830144 @CGN 1 1 2441 @runat_04122002_141353_2264 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMECOUT=120
-WARN_TIMECOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_htc.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: gb_gss.*
18: em_gss_hum.*
19: em_gss_inv.*
20: em_gss_pln.*
21: em_gss_vrt.*
22: em_gss_fun.*
23: em_gss_mam.*
24: em_gss_mus.*
25: em_gss_other.*
26: em_gss_pro.*
27: em_gss_rod.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1500.5	58.2	1036	13	BM560774	BM560774 AGENCOURT
2	1465	56.8	1086	14	BM906573	BM906573 AGENCOURT
3	1376.5	53.4	883	14	BQ929862	BQ929862 AGENCOURT
4	1352	52.4	863	14	BQ322049	BQ322049 AGENCOURT
5	1289	50.0	867	14	BQ431917	BQ431917 AGENCOURT
6	1280.5	49.6	847	13	BI915996	BI915996 603184416
7	1280	49.6	892	14	BQ644850	BQ644850 AGENCOURT
8	1266.5	49.1	991	14	BQ645086	BQ645086 AGENCOURT
9	1212.5	47.0	784	14	BQ442448	BQ442448 UI-M-EVO-
10	1188.5	46.1	917	14	BQ917606	BQ917606 AGENCOURT
11	1166	45.2	815	13	BI181306	BI181306 UNL-P-FN-
12	1165.5	45.2	910	12	BG424017	BG424017 602447479
13	1150	44.6	714	14	BQ179703	BQ179703 UI-M-EWO-
14	1135	44.0	690	10	AW247232	AW247232 2820685.5
15	1129	43.8	702	12	BG715206	BG715206 602675668
16	1106	42.9	840	13	BI918524	BI918524 603176709
17	1094	42.4	691	12	BG686343	BG686343 602638243
18	1072	41.6	974	12	BE989567	BE989567 601681494
19	1049	40.7	677	13	BI101409	BI101409 602887175
20	1034.5	40.1	959	12	BG250255	BG250255 602362379
21	1034	40.1	1062	12	BE895167	BE895167 601436096
22	1026	39.8	1111	10	BE311720	BE311720 601143309
23	1006	39.0	777	13	BI181203	BI181203 UNL-P-FN-
24	1003	38.9	624	10	AW476560	AW476560 ug77g04.Y
25	992	38.4	604	13	BI562276	BI562276 60355406
26	980.5	38.0	705	10	BG623528	BG623528 BB623528
27	979.5	38.0	700	9	AJ446923	AJ446923 AJ446923
28	976	37.8	634	13	BI067793	BI067793 P9fln.pk0
29	955	37.0	614	13	BG913752	BG913752 602810906
30	945.5	36.6	742	13	BI18875	BI18875 603180927
31	945	36.6	614	12	BF087349	BF087349 QV2-HT054
32	921	35.7	789	13	BG913219	BG913219 602811885
33	911	35.3	546	10	BE257942	BE257942 601109776
34	911	35.3	594	9	AA726609	AA726609 vU93D06.f
35	907	35.2	843	12	BG425637	BG425637 602452939
36	889	34.5	625	9	AL585741	AL585741 AL585741
37	878.5	34.1	818	12	BE746542	BE746542 601580107
38	875	33.9	603	12	BG713339	BG713339 P9fln.pk0
39	872	33.8	664	12	BG820485	BG820485 602782796
40	871	33.8	1159	14	BQ229819	BQ229819 AGENCOURT
41	865	33.5	782	13	BI253992	BI253992 602975130
42	862.5	33.4	527	9	AL118967	AL118967 DKFZP761E
43	838	32.5	635	9	AL783587	AL783587 AL783587
44	826.5	32.0	767	13	BI252061	BI252061 602952724
45	824	31.9	599	13	BI401530	BI401530 MI-P-CP0-

ALIGNMENTS

RESULT 1
BM560774 1036 bp mRNA linear EST 20-FEB-2002
LOCUS AGENCOURT_6566220 NIH_MGC_67 Homo sapiens cDNA clone IMAGE:5550695
DEFINITION 5', mRNA sequence.
ACCESSION BM560774
VERSION BM560774.1 GI:18805455
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1036)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL2263 row: d column: 24
 High quality sequence stop: 685.

FEATURES
 source
 1. 1036
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:550695"
 /clone_1b="NIH_MGC_67"
 /tissue_type="retinoblastoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo
 Average insert size 1.75 kb. Library constructed by Life
 Technologies."

BASE COUNT 224 a 285 c 324 g 193 t 10 others
 ORIGIN

Alignment Scores:

Pred. No.: 4.48e-143 Length: 1036
 Score: 1500.50 Matches: 316
 Percent Similarity: 92.77% Conservative: 5
 Best Local Similarity: 91.33% Mismatches: 20
 Query Match: 58.16% Indels: 8
 DB: 13 Gaps: 1

US-09-830-144-4 (1-504) x BM560774 (1-1036)

QY 5 ArgArgSerLeuLeuGlnSerGlnGlnInProSerTyrThrAspAspLeuProLeuCys 24
 Db 6 AGAGAGGAGCTTGTGCGAGATGAGCAGCAGCAGCTGGACAGATACCTGCTCTGCG 65
 QY 25 HisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLeuSerGlyThr 44
 Db 66 CACCTCTCTGGGGTGGCTAGCCTCCACCGCAGCTACTCTGCTGATGGCAAGGCACT 125
 QY 45 GluSerHisProProGluAspSerTyrPheLeuLysPheArgSerGluAsnAsnCysPheLeu 64
 Db 126 GAGAGCCACCCGCGCAGAGCAGACTGCTCAAGTTCAGAGGTGAGCAACAACCTGCTCTG 185
 QY 65 TyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeu 84
 Db 186 TATGGGGTCTTCAACGGCTATGATGCAACCGAGTACCAACTCTGGGCCCAAGCGGCTG 245
 QY 85 SerAlaGluLeuLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGluAlaAspValArgArg 104
 Db 246 TCCGCGAGAGCTCTGCTGGGCCAGCTGAATCCGAGCAGCCAGGCCCATGTGCGCGCT 305
 QY 105 ValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSerIleAspAspAla 124
 Db 306 GTGCTGCTGAGGCTTCCATGTGTGTGAGAGAGAGCTTCCTGGAGTCCATTGACGAGCGCC 365
 QY 125 LeuAlaGluValAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeu 144
 Db 366 TTGGCTGAGAGGCAAGCCCTCCAGTCGCAATTGCCAGAGGAGATCCCTCAGACCAAGCTG 425
 QY 145 ProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArgGlnLeuSerGly 164
 Db 426 CTTCTCTCAGATACGAAGATCTTGGAGAGACTCAAGACCTTAGAGAGGGAAATTTGCGGA 485
 QY 165 GlyAlaMetAlaValAlaValAlaLeuLeuAsnAsnLysLeuTyrValAlaAsnValGly 184
 Db 486 GGGGCCATGGCGCTTGTGGGGTCTTCTCAACAACAAGCTCACTGCGCCATGTGCGGT 545

QY 185 ThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsn 204
 Db 546 ACAACCGTGACCTTTATATGCAATGCACAGTGAATGGTTGAGAGGTGACACAGCTGAAC 605
 QY 205 ValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAsp 224
 Db 606 GTGACACACACACAGAGAAAGAGAGAGACTCTTCGCTCTTCGACAGCTGGCTTGAT 665
 QY 225 AlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnLeuSerThrArgArgIleGly 244
 Db 666 GCTGAAAGATCAAGCAGGCGGGGATATCTGTGGCGAGAGACCCGCGGATCGGG 725
 QY 245 AspTyrLysValLysTyrGlyTyrThrAspIleAspLeuSerAlaAlaLysSerLys 264
 Db 726 GATTACAGGTTAA-TATGGCTACACGAGATTAAGCTTCTCAGCGCTGCCAAGTCCAA 784
 QY 265 ProIleIleAlaGluProGluIleHisGlyAlaGlnPro-LeuAspGlyValThrGlyPh 284
 Db 785 CCAATCATCGCAGAGCCAGAAATCCATGAGGAGCAGCCGCTGGATGGGTGACGGGCTT 844
 QY 284 eLeuValLeu-MetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyProGlyG 304
 Db 845 CTGTGTCTGATTGTGNAGGGGTGTACAAAGCCCTTACAGGAGCCCATGGGCTGGGC 904
 QY 304 InAlaAsnGlnGluIle-AlaAlaMetIleAspThrGluPheAlaLysGlnThrSerLeu 323
 Db 905 AGGCCACCCCGACAGATCTCGATGANNATTCACACTGATGGCCAG-CGACCTCTCCTG 963
 QY 324 AspAlaValAlaGlnAlaValAlaValAspArgVal-----LysArgIleHisSerAspThr 341
 Db 964 GACCAATGGGCCCCAGNCC-GTCCNTGAAACCGGGNTGAGAGCATCCACAGAGACACC 1022
 QY 342 PheAlaSerGly 345
 Db 1023 TTCCCAACCGGG 1034

RESULT 2

BM906573

LOCUS

DEFINITION BM906573 1056 bp mRNA linear EST 12-MAR-2002

AGENCOURT_6621629 NIH_MGC_125 Homo sapiens CDNA clone IMAGE:572376

ACCESSION

BM906573

VERSION

BM906573.1 GI:19356952

KEYWORDS

EST.

SOURCE

ORGANISM

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

NIH-MGC http://mgs.nci.nih.gov/

1 (bases 1 to 1056)

CONTACT: Robert Strausberg, Ph.D.

UNPUBLISHED (1999)

COMMENT

JOURNAL

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL2708 row: f column: 09

High quality sequence stop: 685.

FEATURES

source

1. 1056

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:572376"

/clone_1b="NIH_MGC_125"

/lab_host="DH10B"

/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;

Site_1: BcoRV (destroyed); Site_2: NotI; RNA source pool

of three ovaries, from females ranging in age from 38 to

49 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 2.1 kb, insert size range 1-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 036."

BASE COUNT 239 a 232 c 329 g 196 t
ORIGIN

Alignment Scores:

Pred. No.: 2,02e-139 Length: 1056
Score: 1465.00 Matches: 309
Percent Similarity: 95.69% Conservative: 2
Best Local Similarity: 95.08% Mismatches: 8
Query Match: 56.78% Indels: 7
DB: 14 Gaps: 0

US-09-830-144-4 (1-504) x BM906573 (1-1056)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnProSerTrpThrAspAsp 20
Db 12 ATGGCGGCGCAGAGGAGGAGCTTCTGCTCAGAGTGAGCAGCAAGCTGGACAGATGAC 71
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 72 CTGCCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTCCCAACCGCAGCTACTCTGCTGAT 131
Qy 41 GlyLysGlyThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 132 GGCAAGGCACTGAGACCCACCGCCAGAGACAGCTGGCTCAAGTTCCAGAGTGAGAAC 191
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 192 AACTGCTTCCTGTATGGGCTCTTCAACGGCTATGATGGCAACCGAGTGACCAACTTCGTG 251
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 252 GCCCAGCGGCTGTCGAGAGCTCTCTGGGCCAGCTGAATCCGAGCACGCCGAGGCC 311
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 312 GATGTGGGGCGTGTGCTGCTGAGGCTTCGATGTGGTGAGAGAGCTTCCTGGAGTCC 371
Qy 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 372 ATTGACAGCCCTTGGCTGAGAAGCAAGCTCCAGTCGCAATTGCCAGAGGAGTCCCT 431
Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 432 CAGCACAGCTGCCTCCTCAGTATCAGAAGATCCTTCAGAGACTCAAGACGTTAGAGAGG 491
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 492 GAAATTCGGAGGGGCCATGGCGGTGTGGCGGTCTCTTCTCAACAACCAAGCTCTACGTC 551
Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 552 GCCAATGTCGGTCAAAACCGTCACATTTTATGCAAAATCGACAGTGCATGGGTTGCAGGTG 611
Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 612 ACACAGCTGAACGTGAGACCAACACAGAGAACGAGGATGAGCTCTTCGCTCTTCGCGAC 671
Qy 221 LeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleLeuCysGlyGlnGluSerThr 240
Db 672 CTGGCTTGGATCTGGAAGATCAGCAGGTGGGATCATCTGTGGGCGAGAGACACC 731
Qy 241 ArgArgIleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAla 260
Db 732 CGCGGATCGGGATTACAG-GTTAAATATGGCTACACGACATTGACCTTCTCAGCGCT 790
Qy 261 AlaLysSerLysProIleAlaGluProGluIleHisGlyAlaGln-ProLeuAspGln 280
Db 261 |||||

Db 791 GCCAAGTCCAAACCAATCATCGCAGAGCCGCAATCCCTGGGCGACAGCCCTCGATGG 850
Qy 280 VVal-ThrGlyPheLeuValLeuMetSerGluGlyLeuTyrLysAla-LeuGluAla 299
Db 851 GGTGAACGGCTTCTGGGCTGATGCCGAGGGGTGTACAAGGCCCTTAAGGCGACGC 910
Qy 300 -HisGlyProGlyGlnAla-AsnGlnGluIleAlaAla-MetIleAspThrGluPheAla 318
Db 911 CAGGGGCCCGGGCAGGCCAAACAGAGAAATGTGTCACCAAGATTGACCTGAATTTGC 970
Qy 319 Lys 319
Db 971 CAA 973
RESULT 3
BO929862
LOCUS BO929862
DEFINITION BO929862 883 bp mRNA linear EST 20-AUG-2002
5', mRNA sequence.
ACCESSION BO929862
VERSION BO929862.1 GI:22344893
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 883)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LICM2648 row: c column: 19
High quality sequence stop: 705.
Location/Qualifiers
1. .883
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6460842"
/clone_lib="NIH MGC_101"
/tissue_type="epidermoid carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site: 1: EcoRI; Site 2:
XhoI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Library constructed by Ling Hong in the
Laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH MGC Library."

FEATURES
source

BASE COUNT 195 a 234 c 276 g 175 t 3 others
ORIGIN
Alignment Scores:
Pred. No.: 1.78e-130 Length: 883
Score: 1376.50 Matches: 274
Percent Similarity: 97.86% Conservative: 0
Best Local Similarity: 97.86% Mismatches: 4
Query Match: 53.35% Indels: 2
DB: 14 Gaps: 1
US-09-830-144-4 (1-504) x BO929862 (1-883)

Qy 7 SerLeuLeuGlnSerGluGlnProSerTrpThrAspLeuProLeuCysHisLeu 26
Db 2 AGCTTGCTGCAGAGTGAGCAGCAAGCTGGACATGACCTGCTCTGCGCACCTC 61

27 SerGIValaIGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSer 46
Db 62 TCTGGGGTGGCTCAGCTCCCAACCGCAGCTACTCTGCTGATGGCAAGGAGCTGAGAGC 121
Qy 47 H1sPpProGluAspSerTyrLeuLysPheArgSerGluAsnAsnCysPheLeuTyrGly 66
Db 122 CACCGCCGAGAGACAGCTGGCTCAAGTTCAAGAGTAGAACAATTGCTTCTGTATGG 181
Qy 67 ValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArgLeuSerAla 86
Db 182 GCTTCAACGGCTATGATGCAACGAGTAGCAACTTCTGTGGCCAGCGGCTGTCCGCA 241
Qy 87 GluLeuLeuLeuGlyGlnLeuAsnAlaGlnHisAlaGlnAlaAspValArgArgValLeu 106
Db 242 GAGCTCTCTGCTGGCCAGCTGAATGCCGAGCGCCGAGCGGATGTGGGCTGTGGTG 301
Qy 107 LeuGlnAlaPheAspValValGluArgSerPheLeuGluSerTyrLeuAspAlaLeuAla 126
Db 302 CTGCAAGGCTTTGATGTGTGAGAGAGAGACTTCTGTGAGTCCATTGACGACCCCTTGGCT 361
Qy 127 GluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeuProPro 146
Db 362 GAGAAGCGAAGCTCCAGTCGCAATTCGACAGAGAGTCCCTCAGACACAGCTGCTCCT 421
Qy 147 GlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGlnIleSerGlyGlyAla 166
Db 422 CAGTATCAGAGATCCTTGAAGACTCAAGACGTTAGAGAGGAAATTTGGGAGGGGCC 481
Qy 167 MetAlaValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGlyThrAsn 186
Db 482 ATGGCGGTGTGGGCTCTTCTCAACAAAGCTTACGTGCGCAATGCTCGTACAAAC 541
Qy 187 ArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAsp 206
Db 542 CGGCACTTTTATGCAAAATCGACAGTGGTTCAGGTGACACAGCTGAACGTGAGC 601
Qy 207 HisThrThrGluAsnGlnAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGly 226
Db 602 CACACCACAGAGAACGAGATGAGCTCTTCCGCTTTCACAGCTGGGCTGTGATCTGA 661
Qy 227 LysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyr 246
Db 662 AAGATCAAGAGGTGGGATCATCTGTGGACAGAGAGACCCCGGCGATTCGGGATTC 721
Qy 247 LysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIle 266
Db 722 AAGCTTAAATATGCTACACAGGACATTGACCTTCTCAGCGGCTGCCAAGTCCANACCAATC 781
Qy 267 IleAlaGluProGluIleHis-GlyAlaGlnProLeuAsp---GlyValThrGlyPhe 284
Db 782 ATGCGACAGCCAGAAATCCATGGGGCCACAGCCCGCTGGAGTGGGGTGAACGGGCTTT 839

RESULT 4
BQ922049 863 bp mRNA linear EST 20-AUG-2002
LOCUS BQ922049
DEFINITION AGENCOURT 8923996 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6467676
5', mRNA sequence.
ACCESSION BQ922049
VERSION BQ922049.1 GI:22337080
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 863)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://image.llnl.gov>
Plate: LLM13993 row: p column: 13
High quality sequence stop: 654.
Location/Qualifiers
1. 863
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:6467676"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/note="Organ: eye; Vector: pCMV-SpOrf6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT 193 a 234 c 264 g 171 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 5,59e-128 Length: 863
Score: 1352.00 Matches: 276
Percent Similarity: 97.59% Conservative: 7
Best Local Similarity: 95.17% Mismatches: 2
Query Match: 52,40% Indels: 6
DB: 14 Gaps: 2

US-09-830-144-4 (1-504) x BQ922049 (1-863)

Qy 107 LeuGlnAlaPheAspValValGluArgSerPheLeuGluSerTyrLeuAspAlaLeuAla 126
Db 3 CTTCAGGCTTTGATGTGTGGAGAGAGCTTCTGAGCTATCAATGATGCTTACCT 62
Qy 127 GluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGlnLeuProPro 146
Db 63 GAGAAGCGAAGCTCCAGTCCAGCTGCGCGAGGGTGTGCCAGACACAGCTGCACT 122
Qy 147 GlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGlnIleSerGlyGlyAla 166
Db 123 CAGTATCAGAGATCCTTGAAGACTCAAGACGTTAGAGAGGAAATTTGGGAGGGGCC 182
Qy 167 MetAlaValAlaValLeuLeuAsnLysLeuTyrValAlaAsnValGlyThrAsn 186
Db 183 ATGGCTGTGGGAGTCTTCTCAACAGCAAGCTTACGTTGCCAATGCTCGGTACGAAC 242
Qy 187 ArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAsp 206
Db 243 AGGCGCTTCTGTGCAAAATCTACAGTGGATGGGTACAGGTTACACAGCTTAACATGAC 302
Qy 207 HisThrThrGluAsnGlnAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGly 226
Db 303 CACACCACAGAGACGAGACGAGCTCTTGGCTTTCGAACTGGGTTTAAACGACAGG 362
Qy 227 LysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyr 246
Db 363 AAGATCAAGAGATGGGCTCATCTGTGACAGAGACGACAGCGGCGATGGGATTC 422
Qy 247 LysValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIle 266
Db 423 AAGGTCAAAATATGGCTACCGCAGCATTGACTGTCAAGCTCCCAAGTCCAAACCATC 482
Qy 267 IleAlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuVal 286
Db 483 ATCGAGAGCCGAAATTCATGGGCAACGCTTGTGATGGCTGTGACGGGCTTCTGTG 542
Qy 287 LeuMetSerGluGlyLeuTyrLysAlaLeuGlnAlaAlaHisGlyProGlyGlnAlaAsn 306
Db 543 CTGATGTAGAGGGGCTGTCAAGAGCCCTGAGGACCCATGAGGCTGTGGCAGGCCAAC 602

```

QY 307 GlnGluIleAlaAlaMetIleAspThrGluPheAlaLysGlnThrSer-LeuAspAlaVa 326
Db 603 CAGAGATGCGCGGATGATTGACACCGAGTTTGCCAAAGCAGACCTCCCTGGATCGGT 662
QY 326 lAlaGlnAlaValValAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGl 346
Db 663 TGCCAGGCTGTGTGAGACCGTGTAAGCGGATCCACAGTGAACCTTTGCCAGTGTGG 722
QY 346 yGluArgAlaArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsnPheGl 366
Db 723 GGAGCGTGCCAAAGTTCTGCCACGGCATGAAGACATGACCTGCTGTGAGGGACITTTGG 782
QY 366 yTyrProLeuGlyClnuMetSerGlnProThrProSerProAlaProAlaAlaGlyGlyAr 386
Db 783 CTATCCATTGGTGAATAGAGCCAGCTACACCGACCCCTGCCCA-----GGGGNCC 836
QY 386 sVal---TyrProVal-SerValPro 393
Db 837 -GTGGGTACCTGTTTCTGTGTGCC 861

RESULT 5
BQ431917
LOCUS BQ431917.1 867 bp mRNA linear EST 24-MAY-2002
DEFINITION AGENCOURT_7909771 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6156400
5', mRNA sequence.
ACCESSION BQ431917
VERSION BQ431917.1 GI:21170993
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 867)
AUTHORS NTH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM13500 row: j column: 17
High quality sequence stop: 631.
FEATURES
Location/Qualifiers
source 1..867
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6156400"
/clone_lib="NIH_MGC_71"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site 1: Nth1;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
BASE COUNT 178 a 280 c 257 g 152 t
ORIGIN
Alignment Scores:
Pred. No.: 1..63e-121 Length: 867
Score: 1289.00 Matches: 269
Percent Similarity: 93.79% Conservative: 3
Best Local Similarity: 92.76% Mismatches: 10
Query Match: 49.96% Indels: 8
DB: 14 Caps: 3

US-09-830-144-4 (1-504) x BQ431917 (1-867)

QY 210 GluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIleLys 229

```

```

Db 2 GAGAACGAGGATGAGCTCTTCGCTCTTTTCGAGCTGGGCTTGGATGCTGGAAGATCAAG 61
QY 230 GlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyrLysValLys 249
Db 62 CAGGTGGGATCATCTGTGGCAGGAGAGACCCCGCGGATCGGGGATTACAAGGTTAAA 121
QY 250 TyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIleAlaGlu 269
Db 122 TATGGCTACAGGACATTTGACCTTCTCAGCGCTGCCAAGTCCAAACCAATCATCGCAG 181
QY 270 ProGluIleHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValLeuMetSer 289
Db 182 CCAGAAATCCATGGGGCAGCCGCTGGATGGGTGACGGCTTCTTGGTGTCTGATGTCG 241
QY 290 GluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyProGlyGlnAlaAsnGlnGluIle 309
Db 242 GAGGGTGTGTAAGGCCCTTAGAGCAGCCCATGGGCTGGGAGGCCAACAGAGATT 301
QY 310 AlaAlaMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAla 329
Db 302 GCTGCGATGATTGACATGAGTTTGCCAAAGCAGACCTTCCCTGGACGAGTGGCCAGGCC 361
QY 330 ValValAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGluArgAla 349
Db 362 GTCTGGAGCCGGTGAAGCGCATCCACAGCAGACCTTTCGCGAGTGTGGGAGCGTGCC 421
QY 350 ArgPheCysProArgHisGluAspMetThrLeuLeuValArgAsnPheGlyTyrProLeu 369
Db 422 AGTTTCTCCCGCCGCGCAGGACATGACCTGTGTAGTGAAGAACTTTGGCTACCCCGTG 481
QY 370 GlyGluMetSerGlnProThrProSerProAlaProAlaAlaGlyGlyArgValTyrPro 389
Db 482 GCGAAATGAGCCAGCCACACCGAGCCCGAGCCCGAGCTGCAGAGGACGAGTGTACCT 541
QY 390 ValSerValProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeuSerLeu 409
Db 542 GTGTCTGTGCCATCTCCAGCGGCCAGAGCACCAGCAAGACGAGCGTGACCTCTCCCTT 601
QY 410 ValMetProSerGlnGlyGlnMetValAsnGlyAlaHisSerAlaSerThrLeuAspGlu 429
Db 602 GTCATGCCCTCCAGGGCCAGATGTCACCGGGCTCACAGTGTCTTCCACCTCGAGCAA 661
QY 430 AlaThrProThrLeuThrAsnGlnSerProThrLeuThrLeuGlnSer-ThrAsnThrHi 449
Db 662 GCCACCCCAACCTTCCACCAACCAAGCCCGACCTTAACCTTGCAGTCCACCAACACGCA 721
QY 449 sThrGlnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 468
Db 722 CACGCAGAGCAGCAGCTCCCGCTCTGACGAGGCGCTCTTCCGCTTCGGGCGCCGCCAC 781
QY 469 SerLeuPro-----ProGlyGluAspGlyArgValGluProTyrVal-AspPhe---Al 485
Db 782 TCGGTCCCGCGCTGCGCCAAAGGCGGCGCTGTTTGAAGCCCTTATGTGGGACTTTTGGC 841
QY 485 aGluPheTyr---ArgLeuTyr 491
Db 842 TGAGTTTTTACCGCCCTCTCG 863

RESULT 6
BQ15996
LOCUS BQ15996
DEFINITION 603184416F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5248730 5',
mRNA sequence.
ACCESSION BQ15996
VERSION BQ15996.1 GI:16179953
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 847)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM11627 row: 0 column: 03
High quality sequence stop: 841.
Location/Qualifiers

FEATURES
source

1..847
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5248730"
/clone_lib="NIH_MGC_121"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: Not1; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb. Insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH MGC Library."
BASE COUNT 195 a 229 c 264 g 159 t
ORIGIN

Alignment Scores:

Pred. No.: 1,17e-120 Length: 847
Score: 1280.50 Matches: 269
Percent Similarity: 95.42% Conservative: 2
Best Local Similarity: 94.72% Mismatches: 6
Query Match: 49.63% Indels: 7
DB: 13 Gaps: 1

US-09-830-144-4 (1-504) x BI915996 (1-847)

QY 4 GlnArgSerLeuLeuGlnGlnProSerTrpThrAspLeuProLeu 23
DB 3 CAGAGAGAGAGCTTGGCTGACAGTGAAGCAAGCTGACAGATGACCTTC 62
QY 24 CysHisLeuSerGlyValGlySerAlaSerAspArgSerTyrSerAlaAspGlyLysGly 43
DB 63 TGCACCTCTCTGGGGAGGCTCAGCTCCACCGAGCTACTCTCTGATGGCAAGGC 122
QY 44 ThrGluSerHisProProGluAspSerTrpLeuLysPheArgSerGluAsnAsnCysPhe 63
DB 123 ACTGAGAGCAACCCGCCAGAGGACAGCTGCTCAAGTTCAGAGGTAGAAACAATCTTC 182
QY 64 LeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheValAlaGlnArg 83
DB 183 CTGATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGCACAACTTCGTGGCCAGCCG 242
QY 84 LeuSerAlaGluLeuLeuLeuGlnGlnLeuAsnAlaGlnHisAlaGlnLysAspValArg 103
DB 243 CTGTCGCCAGAGCTCCGTGGGCGACAGTCAATGCGAGCACCCCGAGCCGATGTCCG 302
QY 104 ArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGlnSerIleAspAsp 123
DB 303 CGGTGTGCTGACAGGCTTCGATGTGTGGAAGAGGCTTCCTGAGTTCATTCAGAC 362
QY 124 AlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValProGlnHisGln 143
DB 363 GCCTTGCTGAGAAAGCAAGCTCCAGTCCCAATTGCGAGAGGAGTCCCTCAGCACCA 422
QY 144 LeuProProGlnTyrGlnLysIleLeuGlnArgLeuLysThrLeuGlnArgGlyLysSer 163
DB 423 CTGCTCTCTCAATCAAGAGATCTTGAGAGACTCAAGCGTTAAGAGAGGAAATTTG 482

QY 164 GlyGlyAlaMetAlaValValAlaValLeuLeuAsnAsnLysLeuTyrValAlaAsnVal 183
DB 483 GAGGGGGCCATGAGCCGTGTGGGGGCTCTTCACACAAAGCTTCATCGGCAATGC 542
QY 184 GlyThrAsnArgAlaLeuLeuGlnLysSerThrValAspGlyLeuGlnAlaThrGlnLeu 203
DB 543 GGTACAAACCGTGACTTTATGCAATGCACGTGATGGGTTCAGAGTGACACAGCTG 602
QY 204 AsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeu 223
DB 603 AACGTGACCAACACACAGAGAAAGAGATGAGCTTCGCTTCGACAGCTGGCTTG 662
QY 224 AspAlaGlyLysIleLysGlnValGly-IleIleCys-GlyGlnGlu-SerThrArgArg 242
DB 663 GATGCTGAAAGATCAAGCGAGTGGCGGATCATCTGTGGGCGAGAGACCCGCGG 722
QY 243 Ile-GlyAspTyrLysValLysTyrGlyTyrThr-AspIleAspLeuLeu-SerAlaAla 261
DB 723 ATCGGGGATTTACAGGTTCAATATGGCTACCGGACATTGACCTTCCAGCGCTGCC 782
QY 262 LysSerLysProIleIleAlaGluPro--GluIleHisGlyAlaGlnProLeuAspGly 280
DB 783 AAGTCAACCAATCATCGACAGAGCCAGAAATCCATGGGCGACAGCGCTGATGG 842

RESULT 7
B0644850
LOCUS
DEFINITION
AGENCOURT 8287994 NIH_MGC_100 Homo sapiens cDNA clone IMAGE:6298124
5', mRNA sequence.
ACCESSION
B0644850
VERSION
B0644850.1 GI:21769022
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 892)
NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabs-r@mail.nih.gov
Tissue Procurement: CGAP (Stanford)
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:
http://image.llnl.gov
Plate: LLM2508 row: k column: 21
High quality sequence stop: 637.
Location/Qualifiers

FEATURES
source

1..892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6298124"
/clone_lib="NIH_MGC_100"
/tissue_type="hepatocellular carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."
BASE COUNT 202 a 255 c 273 g 160 t 2 others
ORIGIN

Alignment Scores:

QY 223 LeuAspAlaGlyLysIleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArg 242
 Db 303 TTGGATGCTGGAAAAGATCAAGACAGGTGGGGATCATCTGTGGGACAGAGACCCGGGG 362
 QY 243 TleGlyAspTyrLysValLysTyrGlyThrAspIleAspLeuLeuSerAlaAlaLys 262
 Db 363 ATCGGGGATTACAGATTAAATATGCTACAGACGACATTGACCTTCTCAGCGCTGCCAAG 422
 QY 263 SerLysProIleIleAlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThr 282
 Db 423 TCCAAACCAATATATCGACAGCCAGAAATCCATGGGGCAGACCGCTGGATGGGGTACG 482
 QY 283 GlyPheLeuValLeuMetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyPro 302
 Db 483 GGCTCTTGCTGCTGATGTCCGAGGGGTTGTACAGAGCCCTAGAGCGACCCATGGGCT 542
 QY 303 GlyGlnAlaAsnGlnGluIleAlaIleMetIleAspThrGluPheAlaLysGlnThrSer 322
 Db 543 GGGCAGGCCAACCCAGAGATTGCTCGATGATTGACACTGAACTTTGCCAAGCAGACCTTC 602
 QY 322 LeuAspAlaValAlaGlnAlaValAlaAspArgValLysArgIleHisSerAspThrPhe 342
 Db 603 CTGGACGCGAGTGGCCCGCCGCTGTGACCGGGTGAAGCGCATCCACAGGACACCTTC 662
 QY 343 AlaSerGly--GlyGluArgAlaArgPheCysProAlaGHisGluAspMetThr--LeuLe 361
 Db 663 NNCAGGGGTGGGGAGCGCTGCCANCTTCTGCCCCGACAGACATGACCTCTGCT 722
 QY 361 ValAlaArgAsnMetGlyTyr--ProLeuGlyGluMetSerGlnProThrProSerProAla 380
 Db 723 AGTGAAGGAATCTTGCTTACCCCTGCCGCGAAATGAGCCACCCACCGAGCGCCGCC 782
 QY 381 ProAlaAla--GlyGlyArgValTyrPro--ValSerValProTyrSer--SerAlaGlns 399
 Db 783 CCANCTTGCAGAGAGACCAATGTACCTCGGGGTCTGTGCATCTCCAGGGGCCCAA 842
 QY 399 eThr--SerLysThrSer--ValThrLeuSerLeuValMetProSerGlnGlyMet-- 417
 Db 843 GCACCCACCAAGACCGAGCGGCGCTCTCCCTTGCCATGCTCCCGCAGGGCCAAAAG 902
 QY 418 -ValAsnGlyAlaHis--SerAlaSerThrLeuAspGluAlaThrProThr--LeuThrA 436
 Db 903 GCCACAGGGGGCTCAAAAGGGGCTTCCCCCGGAAAGAAACCCCCCACCCTTACCA 962
 QY 436 sngInserProThrLeu 441
 Db 963 AACAAAGGCCCAACTT 979
 RESULT 9
 BQ442448 784 bp mRNA linear EST 29-MAY-2002
 LOCUS UI-M-EV0-dxe-g-20-0-UI.r1 NIH BMAP_EV0 Mus musculus cDNA clone
 DEFINITION IMAGE:5707051 5', mRNA sequence.
 ACCESSION BQ442448
 VERSION BQ442448.1 GI:21245560
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 784)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>

This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)
 Seq primer: pyx-5,
 Location/Qualifiers
 FEATURES
 source 1..784
 /organism="Mus musculus"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5707051"
 /clone_1ib="NIH BMAP_EV0"
 /tissue_type="whole brain"
 /dev_stage="embryo 15.5 dpc"
 /lab_host="MD10B (T1 phage resistant)"
 /note="Organ: brain; Vector: pyx-asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according to
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured mRNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with an
 oligo-dT primer containing a Not I site. Double stranded
 cDNA was size selected according to mRNA size fraction,
 ligated with EcoR I adaptor, digested with Not I, and then
 cloned directionally into pyx-asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is GTCGCTGGA. This library was created for the
 University of Iowa Mouse Brain Molecular Anatomy Project
 (BMAP): 'Gene Discovery in the Developing Mouse Nervous
 System', supported by National Institutes of Mental Health
 (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 180 a 228 c 223 g 150 t 3 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 9,64e-114 Length: 784
 Score: 1212.50 Matches: 245
 Percent Similarity: 96.18% Conservative: 7
 Best Local Similarity: 93.51% Mismatches: 8
 Query Match: 47,00% Indels: 4
 DB: 14 Gaps: 1

US-09-830-144-4 (1-504) x BQ442448 (1-784)
 QY 188 AlaLeuLeuCysLysSerThrValAspGlyLeuGlnValThrGlnLeuAsnValAspHis 207
 Db 3 GCCCTTGTGCAAAATTTACATGATGATGGTTACAGCTTACACAGCTAAACATGACAC 62
 QY 208 ThrThrGluAsnGluAspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLys 227
 Db 63 ACCACGAGAACGAGACGAGCAGCTCTTCGGCTTCCCACTGGGTTAGACGACGAGGAAG 122
 QY 228 TleLysGlnValGlyIleIleCysGlyGlnGluSerThrArgArgIleGlyAspTyrLys 247
 Db 123 ATCAAGCAGATGGCGCATCTGTGACAGAGACCCAGGGGATGGGGATTACAG 182
 QY 248 ValLysTyrGlyTyrThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIle 267
 Db 183 GTCAAAATATGCTACACCGACATTTGACTCTGACGCTCCCAAGTCCAAACCAATCATC 242
 QY 268 AlaGluProGluIleHisGlyAlaGlnProLeuAspGlyValThrGlyPheLeuValLeu 287
 Db 243 GCAAGCGCGAAATTCATGGCGGACAGCTCTGTGATGCGTGAAGGGCTTCGTGCTG 302
 QY 288 MetSerGluGlyLeuTyrLysAlaLeuGluAlaAlaHisGlyProGlyGlnAlaGln 307
 Db 303 ATGTCAAGGGGCTGTCAAGGGCTGTGAGCAGCCCATGGGCTGGCGCAGCCAACAG 362
 QY 308 GluIleAlaIleMetIleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAla 327
 Db 363 GAGATTGCCCGATGATTGACACCGAATTGGCCAGAGACCTCCCTGATGGGTTGCC 422
 QY 328 GlnAlaValAlaAspArgValLysArgIleHisSerAspThrPheAlaSerGlyGlyGlu 347
 Db 423 CAGCTGTGTGAGACCGGTGTAAAGCGGATCCACAGTACACCTTTGCGAATGGGGAG 482

```

Qy 348 ArgAlaArgPheCysProAraHisGluAspMetThrLeuLeuValArgAsnPheGlyTyr 367
Db 483 CGTGCACAAAGTTCTGTGCCACGGCATGAACACATGACCTTGCTGGTGAGCAACTTTGGCTAT 542

Qy 368 ProLeuGlyGluMetSerGlnProThrProSerProAlaProAlaGlyGlyVal 387
Db 543 CCATTGGGTGAATGAGGACCGCTTACACCGACCCCTTGCCCCA-----GGGGGCGGTGTG 596

Qy 388 TyrProValSerValProTyrSerSerAlaGlnSerThrSerLysThrSerValThrLeu 407
Db 597 TACCCTGTTTCTGTGCCCTACTCAAGTGCCCGAGAGCACCAGACAGACTAGTGTGACTCTG 656

Qy 408 SerLeuValMetProSerGlnGlyGlnMetValAsnGlyAlaHisSerAlaSerThrLeu 427
Db 657 TCNCTCGTCATGCTTCTCAGGGNCAGATGCTCAACGGCTCTCACAGTGCCTCCACTCTG 716

Qy 428 AspGluAlaThrProThrLeuThrAsnGlnSerProThrThrLeuThrLeuGlnSerThrAsn 447
Db 717 GACAAAGCAC--TCCACACTCACTAACAGAG-CCCACTCTGACCTGCAGTCCACANCA 773

Qy 448 ThrHis 449
Db 774 GCACAC 779

RESULT 10
BQ917606
LOCUS
DEFINITION
AGENCOURT 8817353 Lupski_sciatic_nerve Homo sapiens cDNA clone
IMAGE:6202841 5', mRNA sequence.
ACCESSION
BQ917606
VERSION
BQ917606.1 GI:22332304
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 917)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-f@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LILNL at:
http://image.lnl.gov
Plate: LHAM13621 row: i column: 18
High quality sequence stop: 691.
FEATURES
Location/Qualifiers
1..917
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6202841"
/clone_lib="Lupski_sciatic_nerve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6 (Life Technologies); Site_1:
NotI; Site_2: SalI; cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGCGTCCG-3', and
5'-GACTAGTCTAGATCGGAGCGCGCCCT(15)-3'. Size selected >
1 kb for average insert length 1.87 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through Life
Technologies."
198 a 259 c 289 g 171 t
BASE COUNT
ORIGIN

```

```

Alignment Scores:
Pred. No.: 3.64e-111 Length: 917
Score: 1188.50 Matches: 250
Percent Similarity: 90.43% Conservative: 5
Best Local Similarity: 88.65% Mismatches: 18
Query Match: 46.07% Indels: 9
DB: 14 Gaps: 2

US-09-830-144-4 (1-504) x BQ917606 (1-917)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAsp 20
Db 62 ATGGCGGCGCAGAGGAGGAGCTTGCTGCAGAGTGAGCAGCAGCCACCAAGCTTGGACATGAC 121

Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAla 40
Db 122 CTGCTCTCTCTGCCACCTCTCTGGGGTTGGCTCAGCTTCAACCGCAGCTACTCTCTCTGAT 181

Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 182 GGCNAGGGCACTGAGAGCCACCCGCGCAGAGGACAGCTGGCTCAAGTTGAGGAGTGAGAAC 241

Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 242 AACTGTCTTCTGTATGGGGTCTTCAACGGCTATGATGCAACCGAGTGACCAACTTCGTG 301

Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 302 GCCCAGCGGCTGTCCGACAGAGCTCTGTGGGCCAGCTGGAATGGCCGACGACGCCGAGGCC 361

Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 362 GATGTGGCGGTGTCTGCTGCGAGGCCCTTCGATGTGTGGAGAGAGGCTTCTCTGGAGTCC 421

Qy 121 IleAspAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 422 ATTGACGACGCTTGGCTGAGAAGGCAAGCTCCAGTCCGAAATGCCAGAGGAGTCCCT 481

Qy 141 GlnHisGlnLeuProProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 482 CAGCACCACTGCTCTCTCAGTATCAGAAGATCTCTTGAGAGACTCAAGACGTTAGAGAGG 541

Qy 161 GluIleSerGlyValAlaMetAlaValValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 542 GAAATTTCCGAGAGGGCCATGSCCGTTGTGGCGTCTTCTTCAACAACAGCTCTACGTC 601

Qy 181 AlaAsnValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 602 GCCAATGTGCGGTACAAACCGTGCACCTTTATGCAAAATCGACAGTGGATGGGTTCGAGTG 661

Qy 201 ThrGlnLeuAsnValAspHisThrThrGluAsnGluAspGluLeuPheArgLeuSerGln 220
Db 662 ACACAGCTGAACGTGGAGCCACACACAGAGAACGAGGATGAGCTCTTCCGCTTTTCGCAG 721

Qy 221 LeuGlyLeuAspAlaGlyLys-IleLysGlnVal-GlyIleIleCys-GlyGlnGlu-Se 239
Db 722 CTGGGCTTGGATGCTGGAAAGAAATCAGCCAGGTGGGGGATCATCTGCGGGGCGAGGAAG 781

Qy 239 rThrArgArg---IleGlyAspTyrLysValLysTyrGlyTyrThrAspIleAspLeuLe 258
Db 782 CACCCGCGGGAATCGGGGATTACCAGGGTTAAATATGGGTACCC-----CG 829

Qy 258 uSerAlaAlaLysSerLysProIleIleAlaGluProGluIleHisGlyAlaGlnPro 277
Db 830 GACATTGGACCCCTTCTTCCCGCTGCCAGTTCACCAACCAATCATCGCGAGGAGCCA 887

RESULT 11
BQ917606/c
LOCUS
DEFINITION
UNL-P-FN-au-g-01-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
UNL-P-FN-au-g-01-0-UNL 3', mRNA sequence.
ACCESSION
BQ917606

```

```

VERSION      BI181306.1  GI:14655715
KEYWORDS     EST.
SOURCE       Pig.
ORGANISM     Sus scrofa
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE    1 (bases 1 to 815)
AUTHORS      Caetano, A.R., Johnson, R.K. and Pomp, D.
TITLE        Generation and sequence characterization of a normalized cDNA
              library from swine ovarian follicles
              Unpublished (2001)
JOURNAL      Contact: Pomp, D.
              Department of Animal Science
              University of Nebraska, Lincoln
              Lincoln, NE 68583-0908, USA
              Tel: 402 472 6416
              Fax: 402 472 6362
              Email: dpomp@unl.edu
COMMENT      Oligo-dT track not found, Not 1 site shown in beginning of sequence
              is likely internal to the message. The following repetitive
              elements were found in this cDNA sequence: 62-112,
              >GC rich/low complexity
              Seq primer: M13 -29
              POLYA=No.

FEATURES
    source
        1..815
            /organism="Sus scrofa"
            /strain="University of Nebraska, Lincoln Swine Selection
            lines"
            /db_xref="taxon:9823"
            /clone="UNL-P-FN-au-g-01-0-UNL"
            /clone_1ib="UNL-P-FN"
            /dev_stage="ADULT"
            /lab_host="DH10B (Life Technologies)"
            /note="Vector: pUT73D-pac (Pharmacia) with a modified
            polylinker. Site_1: Not 1; Site_2: Eco RI; The UNL-P-FN
            library is a normalized library representing porcine
            ovarian follicles, ranging between 2.0 to 10.0 mm in
            diameter, collected during 7 days of the follicular phase
            of the pig estrous cycle. This library was derived from
            the library UNL-P-F2. The tag is a string of 5-6
            nucleotides present between the Not 1 site and the
            oligo-dT track. The library was constructed as described
            by Bonaldo, Lennon and Soares, Genome Research 6: 791-806
            , 1996.
        TAG_SEQ=None found"
BASE COUNT   109 a 266 c 306 g 134 t
ORIGIN
Alignment Scores:
Pred. No.:    6,05e-109      Length:    815
Score:        1166.00       Matches:   227
Percent Similarity: 98.72%   Conservative: 4
Best Local Similarity: 45.01% Mismatches: 2
Query Match:  45.19%       Indels:      1
DB:           13           Gaps:        0
US-09-830-144-4 (1-504) x BI181306 (1-815)
QY 272 ILHHISGLVLAAGLInProleuAspGIValThnGlyPheluValLeuMeSerGIuGI 291
DB 814 ATCCACGGTGCACAGCCCGCGGAGCGGGTGCCTTCCTGCTGATGTCGAGGGG 755
QY 292 LeuTyTyLeaLeuGluAaIaHISGLYProGIgVlnAla-AsnGLnGIuIleHlaI 311
DB 754 CTITACAAAGCCCTCGAGAGCGACCGGCTCGGGCAGGCCAAACAGAGATCGCAC 695
QY 311 aMeTIIeAsPThGluPheAlaIaLySGInThrSerLeuAspAlaValAlaGlnAlaValVa 331
DB 694 CATGATCGACACAGAGATTGTCCAGACACTCCCTGGATGCGGTGGCCAGGCGCTGT 635
QY 331 laaPaIValVaIaArgIIeHISerAspThrPheAlaSerGIyGIuArgAlaIarph 351

```

```

DB 634 GGACCGGTTGAAGGATCCACAGTGACACTTCGCCAGCGGGGAGCGTCCAGTT 575
QY 351 eCySProArgHISGLIAspMeThrLeuLeuValArgAsnPhelIYTrProLeuGI 371
DB 574 CTGCCACGGCAGAGAGACATGACCTGCTGGTGGCGAATTCGCTACCTCGGCGCA 515
QY 371 uMeSerGIInProThrProSerProAlaProAlaIaGIyGIuArgValTYTrProValSe 391
DB 514 GCTTAGCCACCCACAGCGGGCCCCACGCCAGGAGAGAGAGCGGTATACCCCTGTTC 455
QY 391 rValProTySerSerAlaGInSerThrSerTyThrSerValThnLeuSerLeuValMe 411
DB 454 TGTGCCACTTCAAGGCCGCCAGACACACAGACAGACAGCTGACCTGCTCTCAT 395
QY 411 tProSerGIInGIyGlnMeValAsnGIyAlaHISerAlaSerThrLeuAspGlnAlaTh 431
DB 394 GCCCTCCAGGCGCAGGTGTCAACGGGGCCACAGTGTCTTCACCTCGAGCAACCCAC 335
QY 431 rProThrLeuThrAsnGlnSerProThrLeuThrLeuGInSerThrAsnThHISThrGI 451
DB 334 GCCACCTTCACCAACAGAGCCCGACCTGACCTTGACGTCCACCAACCCACCA 275
QY 451 nSerSerSerSerSerSerAspGIyGlnLeuPheArgSerArgProAlaHISerLeuPr 471
DB 274 GAGCAGACGCTCCAGCTCCACGGGGGCTCTTCGCTCCGCCACACCTCGCTCCC 215
QY 471 oProGIyGIuAspGIyArgValGIuProTYTrValaAspPheAlaGluPheTYrArgLeuTr 491
DB 214 GCCCGGAGAGATGCGCGCGTGGAGCCCTGACGCTTCGAGATTCCGAGATTACCGCTGTG 155
QY 491 pSerValAspHISGLIyGlnGlnSerValValThrAlaPro 504
DB 154 GAGGTGACACCGCGCGAGAGATGTGGAGCGGGCGG 115
RESULT 12
BG424017 910 bp mRNA linear EST 14-MAR-2001
BG424017
LOCUS 602447479F1 NIH_MGC_14 Homo sapiens cDNA clone IMAGE:4586298 5',
DEFINITION mRNA sequence.
ACCESSION BG424017
VERSION BG424017.1 GI:13330523
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE NIH-MGC http://mgi.nci.nih.gov/.
          1 (bases 1 to 910)
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov
          Tissue Procurement: DCTD/DRP
          cDNA Library Preparation: Ling Hong/Rubin Laboratory
          cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
          DNA Sequencing by: Incyte Genomics, Inc.
          Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLCM314 row: m column: 19
          High quality sequence stop: 692.
FEATURES
    source
        1..910
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /clone="IMAGE:4586298"
            /clone_1ib="NIH_MGC_14"
            /tissue_type="renal cell adenocarcinoma"
            /lab_host="DH10B (phage-resistant)"
            /note="Organ: Kidney; Vector: pOTB7; Site_1: XhoI; Site_2:
            EcoRI; cDNA made by oligo-dT priming. Directionally
            cloned into EcoRI/XhoI sites using the following 5'
            adaptor: GGCAAGAG(G). Size-selected >500bp for average

```

insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 197 a 241 c 297 g 175 t

ORIGIN

Alignment Scores:

Pred. No.: 8.21e-109 Length: 910
Score: 1165.50 Matches: 256
Percent Similarity: 92.81% Conservative: 2
Best Local Similarity: 92.09% Mismatches: 15
Query Match: 45.17% Indels: 8
DB: 12 Gaps: 1

US-09-830-144-4 (1-504) x BG424017 (1-910)

Qy 1 MetAlaAlaGlnArgSerLeuLeuGlnSerGluGlnGlnProSerTrpThrAspAsp 20
Db 9 ATGGCGGCGCAGAGAGAGGCTTCTGCAGAGTCAGCAGCAGCAGCCAACTGGACAGATGAC 68
Qy 21 LeuProLeuCysHisLeuSerGlyValGlySerAlaSerAsnArgSerTyrSerAlaAsp 40
Db 69 CTGCTCTCTCCACCTCTCTGGGGTTGGCTCAGCTCCACCGCAGCTACTCTGTGAT 128
Qy 41 GlyLysGlyThrGluSerHisProGluAspSerTrpLeuLysPheArgSerGluAsn 60
Db 129 GGCAGGGCACTGAGAGCCACCCGAGAGCAGCTGGCTCAAGTTCAGAGGTGAGAAC 188
Qy 61 AsnCysPheLeuTyrGlyValPheAsnGlyTyrAspGlyAsnArgValThrAsnPheVal 80
Db 189 AACTGCTTCCTGTATGGGGTCTTCAACGGCTATGATGGCAACCGAGTGACCACTTCGTG 248
Qy 81 AlaGlnArgLeuSerAlaGluLeuLeuGlyGlnLeuAsnAlaGluHisAlaGluAla 100
Db 249 CCCAGCGCTGTTCGAGAGCTCTGCTGGCCAGCTGAATGCCGAGCAGCCGAGGCC 308
Qy 101 AspValArgArgValLeuLeuGlnAlaPheAspValValGluArgSerPheLeuGluSer 120
Db 309 GATGTGGCGCTGTGCTCAGCGCTTCGATGTGTGGAGAGGAGCTTCCTGGAGTCC 368
Qy 121 IleAspAlaLeuAlaGluLysAlaSerLeuGlnSerGlnLeuProGluGlyValPro 140
Db 369 ATTACGACGCTTGGCTGAGAAGCAAGCTCCAGTTCGAATTCGACAGGGAGTCCCT 428
Qy 141 GlnHisGlnLeuProGlnTyrGlnLysIleLeuGluArgLeuLysThrLeuGluArg 160
Db 429 CAGCACCACTGCTCTCTCAGTATCAGAAGATCTTGGAGAGACTCAAGAGCTTAGAGAG 488
Qy 161 GluIleSerGlyGlyAlaMetAlaValAlaValLeuLeuAsnAsnLysLeuTyrVal 180
Db 489 GAAATTCGGAGGGCCATGGCGCTTGTGGCGTCTCTCAACAACAAGCTCTAGCTC 548
Qy 181 AlasenValGlyThrAsnArgAlaLeuLeuCysLysSerThrValAspGlyLeuGlnVal 200
Db 549 GC-AATGTTCGTACAAACCGTCACCTTTATGCAATTCGACAGTGGATGGTTCGAGGTG 607
Qy 201 ThrGlnLeuAsnVal-AspHisThrThrGluAsnGluAspGluLeuPhe-ArgLeu-Ser 219
Db 608 ACACAGCTGAACGTTGGACCAACACACAGAGAACGAGGATGAGCTCTTCCCGCTTTTCG 667
Qy 220 GlnLeuGlyLeuAspAlaGlyLysIleLysGlnValGlyIleIleCys-GlyGlnGluSe 239
Db 668 CAGCTGGGTCGGATGCTGGAAAGATCAGCAGGTGGGGATCATCTGTGGGGCAGGAGAG 727
Qy 239 rThrArgIleGlyAspTyrLysValLysTyrGlyThrThrAspIleAspLeuLeuSe 259
Db 728 CACCGGGGGATCGGGATTA-ACCGTTAACTATGGGTTACCGGA-ATTGACTTCTTAGG 785
Qy 259 rAlaAlaLysSerLysProIleIleAlaGluProGluIleHisGly 274
Db 786 GGGTGCC---AGTCAACAACATCATCGGAGAGGCCCAAAATCTGGGGG 828

RESULT 13
BQ179703
LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BQ179703 714 bp mRNA linear EST 30-APR-2002
UI-M-EWO-Bwu-a-20-0-UI_r1 NIH BMAP_EWO Mus musculus cDNA clone
IMAGE:5703067 5', mRNA sequence.

BQ179703
BQ179703.1 GI:20355195
EST.
house mouse.
Mus musculus.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 714)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LUNL at:
http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

Location/Qualifiers

1..714

/organism="Mus musculus"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE:5703067"

/clone_lib="NIH BMAP_EWO"

/tissue_type="whole brain"

/dev_stage="embryo 15.5 dpc"

/lab_host="DH10B (T1 phage resistant)"

/note="Organ: brain; Vector: pYX-Asc; Site 1: EcoR I;

Site 2: Not I; The library was constructed according to

Bonaldi, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured mRNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with an

oligo-dT primer containing a Not I site. Double stranded

cDNA was size selected according to mRNA size fraction,

ligated with EcoR I adaptor, digested with Not I, and then

cloned directionally into pYX-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

, is GTGCGTGAA. This library was created for the

University of Iowa Mouse Brain Molecular Anatomy Project

(BMAP): 'Gene Discovery in the Developing Mouse Nervous

System', supported by National Institutes of Mental Health

(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 167 a 186 c 213 g 146 t

ORIGIN

Alignment Scores:

Pred. No.: 2.12e-107 Length: 714

Score: 1150.00 Matches: 228

Percent Similarity: 97.90% Conservative: 5

Best Local Similarity: 95.80% Mismatches: 4

Query Match: 44.57% Indels: 2

DB: 14 Gaps: 0

US-09-830-144-4 (1-504) x BQ179703 (1-714)

Qy 9 LeuGlnSerGluGlnGlnProSerTrpThrAspLeuProLeuCysHisLeuSerGly 28

Db 2 CTCGAGGTGACGACGCCAAGCTGGACAGATGACCTGTCACCTCTCTCGA 61

Qy 29 ValGlySerAlaSerAsnArgSerTyrSerAlaAspGlyLysGlyThrGluSerHisPro 48

Db 62 GTTGTGTCAGCGTCCAAACCGCAGCTACTCTGCTGATGCAAGGGCACTGAGAGCCACCT 121

OY	49	ProgluaspserTrrpleulysPheargserGluasnaAsyPheleulnyrGlValaPhe	68
Db	122	CCGAGGACCAACTGGCTTTAAgTTCGAAGGTGAAAAATTAAGTGCCTTGACGGGGCTTC	181
OY	69	AasnGLYTrASPGLVAsnARgVALThrasnpheVALAAGlnARgLeusERLAcluleu	88
Db	182	AATGGCTATGATGGCAACC GGCGTGACAATTGtggcgcCAGAAGGCTTCTGCAGAGCTC	241
OY	89	LeuleuGLYGLInLeuasnaLacLUHIsalagLUAlaSPVALArgArGVAlleuleuGLIn	108
Db	242	CTGCTGGGCGACACTCAACACTGAGACAACCTGAGGCTGAATGTCGACGGGTCTCTGCTGCAG	301
OY	109	AlaPheASPVALValGUlnARgSerPheleuGLUsERTlleasPASALaleuNlaglULys	128
Db	302	GCTTTGATGTGTGTGAAGGAGCTTCCTGGAGTCTATGATGAAGGCCCTTACCTGAGAAA	361
OY	129	AlaSERleuGLInSERCInleuProGIUGLYValPROGLNHISGLInleuProGINtyr	148
Db	362	GCAACCTTCACATCCAGCTGCGGAGGGTGTGCCACACACAGCTGCCACTCAGTAC	421
OY	149	GLInYSllleuGLInARgLEuLSThrIleuGLInARGLInILeserGLYAlametalA	168
Db	422	CAGAAGATCTTGAAGAGACTCCAAGCACCTGAGAGGAGAAATTTCCGGAGACACCATAGCT	481
OY	169	VAlVLAlaValleuLeuasnsnLysleuTYrVALAlaASnVALGYThrasnARgALA	188
Db	482	GTCTGTGGCAGTCTCTTCAACAGCAAAGCTTACGTGCCAAATGTGGTACGAACAGGGCC	541
OY	189	LeuleuCYLSysSERThrVALASpGLyeuGLInVALThrgInLeuasnVALAsphISThr	208
Db	542	CTTCTGTGCAATCTCAAGTGATGGGTTCACAGGTTACACAGCTAAACATGACCAACACC	601
OY	209	Thr-GLInsnGLUASPGLInleuPheARGleuSERGLInleuGLYleuASPALIGLYLSI	228
Db	602	ACCNGAGAACGAGGACGAGACTCTTTCGGCTTTCGCAACGTGGTTTAAAGACGAGNAAGAT	661
OY	228	eLYSGInVALGYlllelleICeYGlyGLInLUSerThraRGAgyILEGLYASP	245
Db	662	CAGACGATGGGCGTCATCTGTGGACAGAGAGCAC-AAGCGGATTTGGGGAT	712
RESULT 14			
LOCUS	AM247232	690 bp	mRNA linear EST 07-JAN-2000
DEFINITION	2820685.5prIME NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2820685 '5'		
ACCESSION	AM247232		
VERSION	AM247232.1		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Euteria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgs.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Other_ESTs: 2820685.3prIME Contact: Robert Strausberg, Ph.D. Email: cgsbpb@mail.nih.gov Tissue Procurement: DCTD/DNP cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL) DNA Sequencing by: Berkeley MGC sequencing project Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www.bio.litnl.gov/btrp/image/image.html Base Calling / Quality Scores: PHRED from University of Washington Genome Center. Vector Trimming: cross match from University of Washington Genome Center PHRAP suite. Poly-T Identification: patmatch.pl from Berkeley Drosophila Genome Project. University of Washington Genome Center http://www.genome.washington.edu Plate: LITCM4 row: N column: 14 High quality sequence stop:513.		

FEATURES		Location/Qualifiers
Source	1. .690	/organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:2820685" /clone_1kb="NIH MGC 7" /tissue_type="small cell carcinoma" /cell_line="MGC3" /lab_host="DH10B (phage-resistant)" /note="Organ: lung; Vector: pOTB1; Site 1: XhoI; Site EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAAGAG (G). Size-selected >500bp for averta insert size 1.8kb. Library constructed by Ling Hong the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)
BASE COUNT	160 a 182 c 217 g 131 t	
ORIGIN		
Alignment Scores:		
Pred. No.:	6.9e-106	Length: 690
Score:	1135.00	Matches: 224
Percent Similarity:	99.13%	Conservative: 3
Best Local Similarity:	97.82%	Mismatches: 2
Query Match:	43.99%	Indels: 0
DB:	10	Gaps: 0
US-09-830-144-4 (1-504) X AM247232 (1-690)		
QY	133 SerGlnLeuProGluGlyValPProGlnHisGlnLeuProProGlnTyrGlnIlySileu	152
Db	4 ACGAGGTGGCAGAGGGGAGTCCCTCACACACAGCTGCTCTCGTACGAAGATCTT	63
QY	153 GIUATGLeuLysThrLeuGLuArgGluileSerGIyGluIleMetAlaValAlaIval	172
Db	64 GAGAGACTCAAGACCTTAGAGAGGGAAATTTGGGAGGGGCGCATGCCGTGTGGCGGTC	123
QY	173 LeuLeuAsnAsnLysLeuTyrValAlaAsnValGIyThrsAsnArgAlaLeuLeuCylys	192
Db	124 CTTCTCAACAAACAAGCTCTACGTCGCCAATGGGTACAAACCGTCGACTTTATGCAA	183
QY	193 SerThrValAspGlyLeuGlnValThGlnLeuAsnValAspHisThrThrGluAsnGlu	212
Db	184 TCGACACGTGATGGGTGGTGCAGGTGACAGCTGAACGTGGCCACACACAGAGACGAG	243
QY	213 AspGluLeuPheArgLeuSerGlnLeuGlyLeuAspAlaGlyLysIleLysGlnValGIy	232
Db	244 GATGAGCTCTTCGCTTCCTCCAGCTGGGCTTGGATGTGGAAAGATCAAGCAGTGGG	303
QY	233 IleIleCySGlyGlnGluSerThrArgArgIleGIyAspTyrLysValIysTyrGIyTyr	252
Db	304 ATCACTGTGGGCGAGAGAGACCCGGGGATTCGGGATTAACAAGTTAATATGGCTAC	363
QY	253 ThrAspIleAspLeuLeuSerAlaAlaLysSerLysProIleIleAlaGluProGluIle	272
Db	364 ACGGACATTGACCTTCTTCAGCGGTGCCAAGTCCAAACCAATCATCGACAGCCAGAAATC	423
QY	273 HisGlyIleArgIleProLeuAspGlyValThGlyPheLeuValLeuMetSerGluGlyLeu	292
Db	424 CATGGGCAACGCCCTGGATGGGATGACGGGGCTTCTTGGTGGTGAATGGAGGGGATG	483
QY	293 TyrLysAlaLeuGluAlaAlaHisGIyProGIyGlnAlaAsnGlnGluIleAlaAlaMet	312
Db	484 TTTCAAGGCCCTTAGAGAGAGCCCATGGGCTGTGAGGCGCAACGAGATTTGGCGATG	543
QY	313 IleAspThrGluPheAlaLysGlnThrSerLeuAspAlaValAlaGlnAlaIvalIvalAsp	332
Db	544 ATTGACACTGATTTGCCAACAACACCTCCCTGGAACCAAGTGGCCCGACCGGTGTGGAC	603
QY	333 ArgValIlyAsnGlyIleHisSerAspThrPheAlaSerGIyGlyGluArgAlaArgPheCys	352
Db	604 CGGGTGAAGCGGATTCAAGGAGACCTTGGCCAGTGGTGGGAGCGTGGCCAAATTGTC	663